On the Two-user Interference Channel with Partial Codebook Knowledge at one Receiver

Alex Dytso, Daniela Tuninetti, and Natasha Devroye

Abstract

In multi-user information theory it is often assumed that every node in the network possesses all codebooks used in the network. This assumption may be impractical in distributed ad-hoc, cognitive or heterogeneous networks. This work considers the two-user Interference Channel with one *Oblivious Receiver* (IC-OR), i.e., one receiver lacks knowledge of the interfering codebook while the other receiver knows both codebooks. The paper asks whether, and if so how much, the channel capacity of the IC-OR is reduced compared to that of the classical IC where both receivers know all codebooks. A novel outer bound is derived and shown to be achievable to within a constant gap for the class of injective semi-deterministic IC-ORs; the gap is shown to be zero for injective fully deterministic IC-ORs. An exact capacity result is shown for the general memoryless IC-OR when the non-oblivious receiver experiences very strong interference. For the linear deterministic IC-OR that models the Gaussian noise channel at high SNR, non i.i.d. Bernoulli(1/2) input bits are shown to achieve points not achievable by i.i.d. Bernoulli(1/2) input bits used in the same achievability same. For the real-valued Gaussian IC-OR the gap is shown to be at most $1/2$ bit per channel use, even though the set of optimal input distributions for the derived outer bound could not be determined. Towards understanding the Gaussian IC-OR, an achievability strategy is evaluated in which the input alphabets at the non-oblivious transmitter are a mixture of discrete and Gaussian random variables, where the cardinality of the discrete part is appropriately chosen as a function of the channel parameters. Surprisingly, as the oblivious receiver

The authors are with the Electrical and Computer Engineering Department of the University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: odytso2, danielat, devroye @uic.edu).

The results in this paper were presented in part at the 2013 IEEE International Symposium on Information Theory [1] and at the 2014 IEEE Information Theory and Applications Workshop [2].

The work of the authors was partially funded by NSF under award number 1017436. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

intuitively should not be able to "jointly decode" the intended and interfering messages (whose codebook is unavailable), it is shown that with this choice of input, the capacity region of the symmetric Gaussian IC-OR is to within $\frac{1}{2}\log(8\pi\mathrm{e}) \approx 3.04$ bits (per channel use per user) of an outer bound for the classical Gaussian IC.

I. Introduction

A classical assumption in multi-user information theory is that each node in the network possesses knowledge of the codebooks used by every other node. However, such an assumption might not be practical in heterogeneous, cognitive, distributed or dynamic networks. For example, in very large ad-hoc networks, where nodes enter and leave at will, it might not be practical for new nodes to learn the codebooks of old nodes and vice-versa. In cognitive radio scenarios, where new cognitive systems coexist with legacy systems, requiring the legacy systems to know the codebooks of the new cognitive systems might not be viable. This motivates the study of networks where each node possesses only a subset of the codebooks used in the network. We will refer to such systems as networks with *partial codebook knowledge* and to nodes with only knowledge of a subset of the codebooks as *oblivious* nodes.

To the best of our knowledge, systems with partial codebook knowledge were first introduced in [3]. In [3] lack of codebook knowledge was modeled by using *codebook indices*, which index the random encoding function that maps the messages to the codewords. If a node has codebook knowledge it knows the index (or instance) of the random encoding function used; else it does not and the codewords essentially look like the symbols were produced in an independent, identically distributed (i.i.d.) fashion from a given distribution. In [4] and [5] this concept of partial codebook knowledge was extended to model *oblivious relays* and capacity results were derived. However, as pointed out in [4, Section III.A] and [5, Remark 5], these capacity regions are "non-computable" in the sense that it is not known how to find the optimal input distribution in general. In particular, the capacity achieving input distribution for the practically relevant Gaussian noise channel remains an open problem.

We make progress on this front by demonstrating that certain rates are achievable for the Gaussian noise interference channel with oblivious receivers (G-IC-OR) through the evaluation of a simplified Han-Kobayashi scheme [6] in which joint decoding of the intended and interfering messages is not required at the oblivious receiver. The major contribution of this work is the realization that Gaussian inputs perform poorly in the proposed achievable region. We therefore propose to use a class of inputs that we termed *mixed inputs*. A mixed input is random variable that is a mixture of a continuous and a discrete part, such as for example a Gaussian random variable and a uniformly distributed random variable on an equally spaced set of discrete points. We then properly design the distribution of the mixed input as a function of the channel parameters.

We are not the first to consider discrete inputs for Gaussian noise channels. In [7] the authors considered the point-to-point power-constrained Gaussian noise channel and derived lower bounds on the achievable rate when the input is contained to be an equally spaced Pulse Amplitude Modulation (PAM) in which each each point is used with equal probability; such an input was shown to be optimal to within 0.41 bits per channel use [7, eq.(9)].

In [8, Theorems 6 and 7], the authors *asymptotically* characterized the optimal input distribution over N masses at high and low SNR, respectively, for a point-to-point power-constrained Gaussian noise channel by assuming that N is not dependent on SNR. For the purpose of this work, these bounds cannot be used, as 1) these bounds are optimized for a specific SNR while we shall need to lower bound the rate achievable by a discrete input at multiple receivers each characterized by a different SNR; 2) we need a *firm* bound that holds at all finite SNR; and 3) we need to properly choose N as a function of SNR, a question posed but left open in [8].

The sub-optimality of Gaussian inputs for Gaussian noise channels has been observed before. Past work on the asynchronous IC [9], [10] showed that non-Gaussian inputs may outperform i.i.d. Gaussian inputs by using local perturbations of an i.i.d. Gaussian input: [9, Lemma 3] considers a fourth order approximation of mutual information, while [10, Theorem 4] uses perturbations in the direction of Hermite polynomials of order larger than three. In both cases the input distribution is assumed to have a density, though [9, Fig. 1] numerically shows the performance of a ternary PAM input as well. For the cases considered in [9], [10], the improvement over i.i.d. Gaussian inputs shows in the decimal digits of the achievable rates; it is hence not clear that perturbed continuous Gaussian inputs as in [9], [10] can actually provide DoF gains over Gaussian inputs (note that a strict DoF gain implies an unbounded rate gain as SNR increases) which we seek in this work.

The major contributions of this work are as follows. The general memoryless IC-OR channel model is introduced in Section II, together with the special class of *injective semi-deterministic*

IC-ORs (ISD-IC-OR) [11] of which the Gaussian noise channel is an example. Then:

1) In Section III we derive a novel outer bound that incorporates this partial codebook knowledge explicitly. In this bound, the single rate bounds are valid for a general memoryless IC-OR while the sum-rate bound is valid for the ISD-IC-OR only.
2) In Section IV we demonstrate a series of capacity and approximate capacity results for various regimes and classes of IC-OR. Specifically: (a) we obtain the capacity region for the general memoryless IC-OR in very strong interference at the non-oblivious receiver, (b) we demonstrate the capacity region to within a constant gap for the ISD-IC-OR, and (c) we show that for the injective fully deterministic IC-OR the gap is zero.
3) Next, we look at the practically relevant G-IC-OR and its corresponding Linear Deterministic Approximation (LDA-IC-OR) in the spirit of [12], which models the G-IC-OR at high SNR. Surprisingly, for the LDA-IC-OR we demonstrate that for a given achievability scheme, i.i.d. Bernoulli(1/2) input bits (known to be optimal for the LDA-IC with full codebook knowledge [13]) are outperformed by other input distributions.
4) For the G-IC-OR, we show that our inner and outer bounds that are to within $1/2$ bit (per channel use per user) of one another. However, similarly to prior work on oblivious models, we are not able to find the set of input distributions that exhaust the outer bound, in other words, we cannot show that i.i.d. Gaussian inputs exhaust the outer bound. Inspired by the results for the LDA-IC-OR, we numerically show that a larger sum-capacity is attainable by using a discrete input at the non-oblivious transmitter than by selecting i.i.d. Gaussian inputs, or using time-division, or treating interference as noise, in the strong interference regime at high SNR. This suggests that the penalty for the lack of codebook knowledge is not as severe as one might initially expect.
5) For the remainder of the paper we consider the G-IC-OR, and demonstrate that even with partial codebook knowledge we are able to achieve to within $\frac{1}{2}\log(8\pi \mathrm{e}) \approx 3.04$ bits per channel use of the *symmetric capacity region* of the G-IC with full codebook knowledge through the use of mixed inputs.[1] The main tool to derive the symmetric capacity to within

[1]The restriction to the symmetric case, i.e., same direct links and same interference links, is just to reduce the number of parameters in our derivations. We strongly believe that an approximate capacity result (to within a constant gap) can be shown for the general asymmetric case, albeit through more tedious computations than those reported here.

a constant gap is the lower bound from [7] on the mutual information achievable by a PAM input on a point-to-point Gaussian noise channel. With this tool, in Section VI we evaluate the achievable rate region presented in Section IV for the G-IC-OR when the non-oblivious transmitter uses a mixed input that comprises a Gaussian component and a PAM component.

6) In past work on networks with oblivious nodes no performance guarantees were provided as the capacity regions could not be evaluated. In Section VII we study the generalized degrees of freedom (gDoF) achievable with mixed inputs. We first show that mixed inputs achieve the gDoF of the classical G-IC, hence implying that there is no loss in performance due to lack of codebooks in a gDoF sense / at high SNR. This is quite surprising considering that the oblivious receiver cannot perform joint decoding of the two messages, which is optimal for the classical G-IC in the strong and very strong interference regimes.
7) Finally, in Section VIII we turn our attention to the finite SNR regime and show that the capacity of the symmetric G-IC-OR is within $\frac{1}{2}\log(8\pi\mathrm{e}) \approx 3.04$ bits per channel use of the outer bound to the capacity region of the classical symmetric G-IC.

We conclude the paper with some final remarks and future directions in Section IX. Some proofs are reported in the Appendix.

II. Channel Model

A. Notation

We adopt the following notation convention:

- Lower case variables are instances of upper case random variables which take on values in calligraphic alphabets.
- Throughout the paper $\log(\cdot)$ denotes logarithms in base 2.
- $[n_1 : n_2]$ is the set of integers from n_1 to $n_2 \geq n_1$.
- Y^j is a vector of length j with components $(Y_1, \ldots, Y_j)$.
- We let $\delta(\cdot)$ denote the the Dirac delta function.
- If A is a random variable (r.v.) we denote its support by $\mathsf{supp}(A)$.
- The symbol $|\cdot|$ may denote different things: $|\mathcal{A}|$ is the cardinality of the set $\mathcal{A}$, $|X|$ is the cardinality of $\mathsf{supp}(X)$ of the r.v. X , or $|x|$ is the absolute value of the real-valued x.
- For $x \in \mathbb{R}$ we let $\lfloor x \rfloor$ denote the largest integer not greater than x.

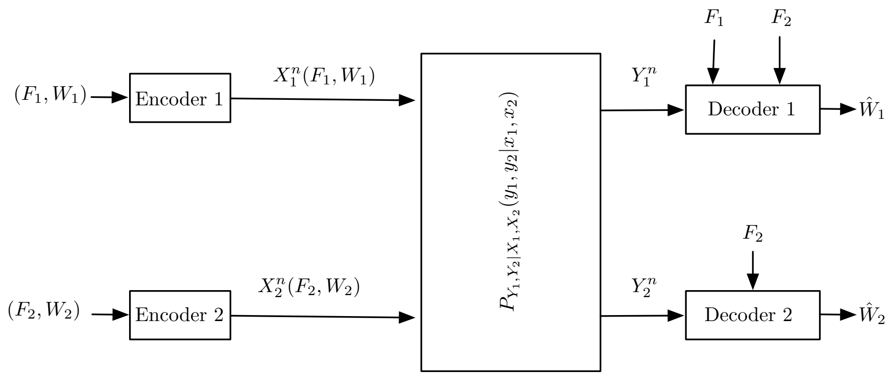


Fig. 1. The IC-OR, where F_1 and F_2 represent codebook indices known to one or both receivers.

- For $x \in \mathbb{R}$ we let $[x]^+ := \max(x, 0)$ and $\log^+(x) := [\log(x)]^+$.
- The functions $\mathsf{I_g}(x)$, $\mathsf{I_d}(n, x)$, $\mathsf{G_d}(x, a)$, and $\mathsf{N_d}(x)$, for $n \in \mathbb{N}$ and $a, x \in \mathbb{R}^+$, are defined as

$$\mathsf{I_g}(x) := \frac{1}{2}\log(1 + x), \tag{1}$$

$$\mathsf{I_d}(n, x) := \left[\frac{1}{2}\log(1 + \min(n^2 - 1, x)) - \frac{1}{2}\log\left(\frac{\pi e}{3}\right)\right]^+, \tag{2}$$

$$\mathsf{N_d}(x) := \left\lfloor \sqrt{1 + x} \right\rfloor, \tag{3}$$

  where the subscript d reminds the reader that discrete inputs are involved, while g that Gaussian inputs are involved.
- $\mathcal{N}(\mu, \sigma^2)$ denotes a real-valued Gaussian r.v. with mean μ and variance σ^2.
- $\mathsf{Unif}([n_1 : n_2])$ denotes the uniform distribution over the set $[n_1 : n_2]$.
- $\mathsf{Bernoulli}(p)$ denotes the Bernoulli distribution with parameter $p \in [0, 1]$.
- $X \sim \mathsf{PAM}(N)$ denotes the uniform distribution over a zero-mean Pulse Amplitude Modulation (PAM) constellation with $|X| = N$ points and unit-energy.
- $\mathsf{co}(\cdot)$ denotes the convex closure operator.

B. General Memoryless IC-OR

An IC-OR consists of the two-user memoryless interference channel $(\mathcal{X}_1, \mathcal{X}_2, P_{Y_1Y_2|X_1X_2}, \mathcal{Y}_1, \mathcal{Y}_2)$ where receiver 2 is oblivious of transmitter 1's codebook. We use the terminology "codebook" to denote a set of codewords and the (one-to-one) mapping of the messages to these codewords. We model lack of codebook knowledge as in [3], where transmitters use randomized encoding functions, which are indexed by a message index and a "codebook index" (F_1 and F_2 in Fig. 1). An oblivious receiver is unaware of the "codebook index" (F_1 is not given to decoder 2 in Fig. 1) and hence does not know how codewords are mapped to messages. The basic modeling assumption is that without the knowledge of the codebook index a codeword looks unstructured. More formally, by extending [4, Definition 2], a $(2^{nR_1}, 2^{nR_2}, n)$ code for the IC-OR with enabled time sharing is a six-tuple $(P_{F_1|Q^n}, \sigma_1^n, \phi_1^n, P_{F_2|Q^n}, \sigma_2^n, \phi_2^n)$, where the distribution $P_{F_i|Q^n}$, $i \in [1:2]$, is over a finite alphabet $\mathcal{F}_i$ conditioned on the time-sharing sequences q^n from some finite alphabet $\mathcal{Q}$, and where the encoders σ_i^n and the decoders ϕ_i^n, are mappings

$$\sigma_1^n : [1:2^{nR_1}] \times [1:|\mathcal{F}_1|] \to \mathcal{X}_1^n,$$
$$\sigma_2^n : [1:2^{nR_2}] \times [1:|\mathcal{F}_2|] \to \mathcal{X}_2^n,$$
$$\phi_1^n : [1:|\mathcal{F}_1|] \times [1:|\mathcal{F}_2|] \times \mathcal{Y}_1^n \to [1:2^{nR_1}],$$
$$\phi_2^n : [1:|\mathcal{F}_2|] \times \mathcal{Y}_2^n \to [1:2^{nR_2}].$$

Moreover, when user 1's codebook index is unknown at decoder 2, the encoder σ_1^n and the distribution $P_{F_1|Q^n}$ must satisfy

$$\begin{aligned}
&\mathbb{P}[X_1^n = x_1^n | Q^n = q^n] \\
&= \sum_{w_1=1}^{2^{nR_1}} \sum_{f_1=1}^{|\mathcal{F}_1|} P_{F_1|Q^n}(f_1|q^n)\ 2^{-nR_1}\ \delta\big(x_1^n - \sigma_1^n(w_1, f_1)\big) \\
&= \prod_{t \in [1:n]} P_{X_1|Q}(x_{1t}|q_t), \qquad (4)
\end{aligned}$$

according to some distribution $P_{X_1|Q}$. In other words, when averaged over the probability of selecting a given codebook and over a uniform distribution on the message set, the transmitted codeword conditioned on any time sharing sequence has an i.i.d. distribution according to some distribution $P_{X_1|Q}$.

A non-negative rate pair (R_1, R_2) is said to be achievable if there exist a sequence of encoding functions $\sigma_1^n(W_1, F_1)$, $\sigma_2^n(W_2, F_2)$, and decoding functions $\phi_1^n(Y_1^n, F_1, F_2)$, $\phi_2^n(Y_2^n, F_2)$, such that the average probability of error satisfies $\max_{i \in [1:2]} \mathbb{P}[\widehat{W}_i \neq W_i] \to 0$ as $n \to +\infty$. The capacity region is defined as the convex closure of all achievable rate pairs (R_1, R_2) [14].

C. Injective Semi-Deterministic IC-OR

For a general memoryless IC-OR, no restrictions are imposed on the transition probability $P_{Y_1Y_2|X_1X_2}$. The ISD-IC-OR is a special IC-OR with transition probability

$$P_{Y_1Y_2|X_1X_2}(y_1, y_2|x_1, x_2) = \sum_{t_1,t_2} P_{T_1|X_1}(t_1|x_1)P_{T_2|X_2}(t_2|x_2) \cdot \delta\big(y_1 - g_1(x_1, t_2)\big)\ \delta\big(y_2 - g_2(x_2, t_1)\big), \tag{5}$$

for some memoryless transition probabilities $P_{T_1|X_1}$ and $P_{T_2|X_2}$, and some deterministic functions $g_1(\cdot, \cdot)$ and $g_2(\cdot, \cdot)$ that are injective when their first argument is held fixed [11]. The ISD property implies that

$$H(Y_1|X_1) = H(T_2) \text{ and } H(Y_2|X_2) = H(T_1), \quad \forall P_{X_1X_2} = P_{X_1}P_{X_2}, \tag{6}$$

or in other words that the T_u is a deterministic function of the pair (Y_u, X_u), $u \in [1 : 2]$. For channels with continuous alphabets, the summation in (5) should be replaced with an integral and the discrete entropies in (6) with the differential entropies.

III. Outer Bounds

In this section we present novel outer bounds for the IC-OR. In particular, we derive the single rate bounds that are valid for a general memoryless IC-OR and a sum-rate bound that is valid for the ISD-IC-OR only.

We begin by proving a property of the output distributions that is key to deriving single-letter expressions in our outer bounds; this property holds for a general memoryless IC-OR.

Proposition 1. *The output of the oblivious decoder has a product distribution conditioned on the signal whose codebook is known, that is,*

$$P_{Y_2^n|X_2^n,F_2}(y_2^n|x_2^n, f_2) = \prod_{i=1}^{n} P_{Y_{2i}|X_{2i}}(y_{2i}|x_{2i}). \tag{7}$$

Proof of Proposition 1: Starting from the joint distribution of Y_2^n, X_1^n conditioned on X_2^n, F_2 we have that

$$\begin{aligned}
&P_{Y_2^n,X_1^n|X_2^n,F_2}(y_2^n, x_1^n|x_2^n, f_2) \\
&\overset{a)}{=} P_{X_1^n}(x_1^n)\prod_{i=1}^n P_{Y_{2i}|X_{1i},X_{2i}}(y_{2i}|x_{1i}, x_{2i}) \\
&\overset{b)}{=} \prod_{i=1}^n P_{X_{1i}}(x_{1i})\prod_{i=1}^n P_{Y_{2i}|X_{1i},X_{2i}}(y_{2i}|x_{1i}, x_{2i}) \\
&\overset{c)}{=} \prod_{i=1}^n P_{Y_{2i},X_{1i}|X_{2i}}(y_{2i}, x_{1i}|x_{2i})
\end{aligned}$$

where the equalities follows from: a) the inputs are independent and the channel is memoryless, b) the assumption that X_1^n has a product distribution if not conditioned on F_1 as in (4), and c) the inputs are independent. By marginalizing with respect to X_1^n yields

$$P_{Y_2^n|X_2^n,F_2}(y_2^n|x_2^n, f_2) = \prod_{i=1}^n \sum_{x_{1i}} P_{Y_{2i},X_{1i}|X_{2i}}(y_{2i}, x_{1i}|x_{2i}) = \prod_{i=1}^n P_{Y_{2i}|X_{2i}}(y_{2i}|x_{2i}),$$

as claimed. ■

The main result of the section is the following upper bound:

Theorem 2. *Any achievable rate pair (R_1, R_2) for the IC-OR must satisfy*

$$R_1 \le I(Y_1; X_1|F_1, F_2, X_2, Q), \quad \textit{(memoryless IC-OR)} \tag{8a}$$

$$R_2 \le I(Y_2; X_2|F_2, Q), \quad \textit{(memoryless IC-OR)} \tag{8b}$$

$$\begin{aligned}
R_1 + R_2 &\le H(Y_1|F_1, F_2, Q) + H(Y_2|F_2, U_2, Q) \\
&\quad - H(T_2|X_2, Q) - H(T_1|Q) \qquad \textit{(memoryless ISD-IC-OR)} \\
&= I(Y_1; X_1, X_2|F_1, F_2, Q) + I(Y_2; X_2|F_2, U_2, Q),
\end{aligned} \tag{8c}$$

for some input distribution that factors as

$$\begin{aligned}
&P_{Q,X_1,X_2,U_2|F_1,F_2}(q, x_1, x_2, u_2|f_1, f_2) \\
&= P_Q(q)P_{X_1|F_1,Q}(x_1|f_1, q)P_{X_2|F_2,Q}(x_2|f_2, q)P_{T_2|X_2}(u_2|x_2),
\end{aligned} \tag{8d}$$

and with $|\mathcal{Q}| \le 2$. We denote the region in (8) *as $\mathcal{R}_{\text{out}}$.*

Proof of Theorem 2: By Fano's inequality $H(W_1|Y_1^n, F_1, F_2) \le n\epsilon_n$ and $H(W_2|Y_2^n, F_2) \le n\epsilon_n$ for some $\epsilon_n \to 0$ as $n \to \infty$.

We begin with the R_1-bound (non-oblivious receiver) in (8a):

$$\begin{aligned} n(R_1 - \epsilon_n) &\overset{a)}{\le} I(W_1; Y_1^n, F_1, F_2) \\ &\overset{b)}{\le} I(W_1; Y_1^n | F_1, F_2, W_2) \\ &\overset{c)}{\le} I(X_1^n; Y_1^n | F_1, F_2, X_2^n) \\ &\overset{d)}{\le} \sum_{i=1}^{n} I(X_{1i}; Y_{1i} | F_1, F_2, X_{2i}), \end{aligned}$$

where the (in)equalities follow from: a) Fano's inequality, b) giving W_2 as side information and using the fact that F_1, F_2, W_1 and W_2 are mutually independent, c) data processing $(F_i, W_i) \to X_i^n \to Y_1^n$, for $i \in [1:2]$, and d) by "conditioning reduces entropy" and the fact that the channel is memoryless.

For the R_2-bound (oblivious receiver) in (8b) we have:

$$\begin{aligned} n(R_2 - \epsilon_n) &\overset{a)}{\le} I(W_2; Y_2^n, F_2) \\ &\overset{b)}{\le} I(W_2; Y_2^n | F_2) \\ &\overset{c)}{\le} I(X_2^n; Y_2^n | F_2) \\ &\overset{d)}{\le} \sum_{i=1}^{n} I(X_{2i}; Y_{2i} | F_2), \end{aligned}$$

where the (in)equalities follow from: a) Fano's inequality, b) the fact that F_2 and W_2 are independent, c) data processing $(F_i, W_i) \to X_i^n \to Y_1^n$, for $i \in [1:2]$, and d) from Lemma.1.

Next, by providing U_2 as side information to receiver 2 (oblivious receiver) similarly to [11], where U_2 is jointly distributed with the inputs according to (8d), we have:

$$\begin{aligned} n(R_1 + R_2 - 2\epsilon_n) &\overset{a)}{\le} I(X_1^n; Y_1^n | F_1, F_2) + I(X_2^n; Y_2^n, U_2^n | F_2) \\ &= H(Y_1^n | F_1, F_2) - H(Y_1^n | F_1, F_2, X_1^n) \\ &+ H(U_2^n | F_2) - H(U_2^n | F_2, X_2^n) \\ &+ H(Y_2^n | F_2, U_2^n) - H(Y_2^n | F_2, X_2^n, U_2^n) \\ &\overset{b)}{=} H(Y_1^n | F_1, F_2) - H(T_2^n | F_1, F_2) \end{aligned}$$

$$
\begin{aligned}
&+ H(U_2^n|F_2) - H(U_2^n|F_2, X_2^n) \\
&+ H(Y_2^n|F_2, U_2^n) - H(T_1^n) \\
&\overset{\text{c)}}{=} H(Y_1^n|F_1, F_2) - H(T_2^n|F_1, F_2) \\
&+ H(T_2^n|F_2) - H(T_2^n|F_2, X_2^n) \\
&+ H(Y_2^n|F_2, U_2^n) - H(T_1^n) \\
&\overset{\text{d)}}{=} H(Y_1^n|F_1, F_2) + H(Y_2^n|F_2, U_2^n) - H(T_2^n|X_2^n) - H(T_1^n) \\
&\overset{\text{e)}}{\leq} \sum_{i=1}^{n} H(Y_{1i}|F_1, F_2) + H(Y_{2i}|F_2, U_{2i}) - H(T_{2i}|F_2, X_{2i}) - H(T_{1i}),
\end{aligned}
$$

where (in)equalities follow from: a) by giving U_2 as side information and by proceeding as done for the single rate bounds up to step labeled "c)", b) by the injective property in (5) and the independence of (X_1^n, T_1^n) and X_2^n, c) by definition of U_2 in (8d) we have $H(U_2^n|F_2) = H(T_2^n|F_2)$, d) by independence of the messages we have $H(T_2^n|F_1, F_2) - H(T_2^n|F_2) = 0$, e) since the channel is memoryless and thus $H(T_2^n|F_2, X_2^n) = H(T_2^n|X_2^n) = \sum_{i=1}^n H(T_{2i}|F_2, X_{2i})$ and since $H(T_1^n)$ can be single-letterized by using Proposition 1.

The introduction of a time-sharing random variable $Q \sim \mathsf{Unif}[1:n]$ yields the bounds in (8). The Fenchel-Eggleston-Caratheodory theorem [15, Chapter 14] guarantees that we may restrict attention to $|Q| \leq 2$ without loss of optimality.

Finally, the equality in (8c) follows from the injective property in (5), the independence of the inputs and the memoryless property of the channel, i.e.,

$$
\begin{aligned}
H(Y_1|F_1, F_2, Q, X_1, X_2) &= H(Y_1|X_1, X_2) = H(T_2|X_1, X_2) = H(T_2|X_2), \\
H(Y_2|F_2, U_2, Q, X_2) &= H(T_1|F_2, U_2, Q, X_2) = H(T_1|Q).
\end{aligned}
$$

This concludes the proof. ■

IV. Capacity Results

In this section we prove that the outer bound in (8) is (approximately) tight in certain regimes or for certain classes of channels. To start, we propose an achievable rate region based on a simplified Han-Kobayashi scheme [6] in which joint decoding of the intended and interfering messages is not required at receiver 2 (the oblivious receiver).

A. Inner Bound

Consider an achievability scheme where encoder 1 transmits using an i.i.d. codebook, while encoder 2, corresponding to the oblivious receiver, rate-splits as in the Han and Kobayashi achievability scheme for the classical IC [6]. It may then be shown that the following rates are achievable, where we have included the codebook index random variables F_1, F_2 to emphasize who knows which codebook:

Proposition 3. *The set of non-negative rate pairs* (R_1, R_2) *satisfying*

$$R_1 \le I(Y_1; X_1 | F_1, F_2, U_2, Q), \tag{9a}$$

$$R_2 \le I(Y_2; X_2 | F_2, Q), \tag{9b}$$

$$R_1 + R_2 \le I(Y_1; X_1, U_2 | F_1, F_2, Q) + I(Y_2; X_2 | F_2, U_2, Q), \tag{9c}$$

is achievable for every input distribution that factorizes as

$$P_{Q,X_1,X_2,U_2|F_1,F_2} = P_Q P_{X_1|F_1,Q} P_{X_2|F_2,Q} P_{U_2|X_2,F_2,Q}. \tag{9d}$$

We denote the region in (9) *as* $\mathcal{R}_{\text{in}}$*, which is achievable for any memoryless IC-OR.*

Proof of Proposition 3: The proof follows by setting the auxiliary r.v. U_1 in the Han and Kobayashi rate region in [14, Section 6.5] to $U_1 = \emptyset$. Note, that this modified version of the Han and Kobayashi scheme employs joint decoding (of desired and undesired messages) only at receiver 1 (the non-oblivious receiver) and hence knowledge of the codebook of transmitter 1 is not needed at receiver 2 (the oblivious receiver). ■

Remark 1. *By comparing the outer bound region* $\mathcal{R}_{\text{out}}$ *in Theorem 2 to the inner bound region* $\mathcal{R}_{\text{in}}$ *in Proposition 3 we notice the following differences: 1) in* (8d) *the side information random variable* U_2 *is distributed as* T_2 *conditioned on* X_2*, while in* (9d) *the auxiliary random variable* U_2 *can have any distribution conditioned on* X_2*; 2) the mutual information terms involving* Y_1 *have* X_2 *in the outer bound, but* U_2 *in the inner bound; and 3) the mutual information terms involving* Y_2 *are the same in both regions.*

B. Capacity in very strong interference at the non-oblivious receiver for the general memoryless IC-OR

In this section we show that under special channel conditions, akin to the very strong interference regime for the classical IC, the outer bound region in Theorem 2 is tight. From now on we drop the codebook indices from Theorem 2 and Proposition 3 in order to simplify the notation.

A general memoryless IC-OR for which

$$I(X_2;Y_2|X_1) \le I(X_2;Y_1), \quad \forall P_{X_1,X_2} = P_{X_1}P_{X_2}, \tag{10}$$

is said to have *very strong interference at the non-oblivious receiver* (receiver 1). Intuitively, when the condition in (10) holds, the non-oblivious receiver should be able to first decode the interfering signal by treating its own signal as noise and then decode its own intended signal free of interference. This should "de-activate" the sum-rate bound in (8c). Next we formalize this intuition.

Theorem 4. *When the condition in* (10) *holds the capacity region of the IC-OR is given by*

$$R_1 \le I(X_1;Y_1|X_2,Q), \tag{11a}$$

$$R_2 \le I(X_2;Y_2|Q), \tag{11b}$$

taken over the union of all input distributions that factor as $P_{Q,X_1,X_2} = P_Q P_{X_1|Q} P_{X_2|Q}$.

Proof of Theorem 4: By dropping the sum-rate outer bound in (8c) we see that the region in (11) is an outer bound for a general memoryless IC-OR. By setting $U_2 = X_2$ in the achievable region in (9), the region

$$R_1 \le I(X_1;Y_1|X_2,Q), \tag{12a}$$

$$R_2 \le I(X_2;Y_2|Q), \tag{12b}$$

$$R_1 + R_2 \le I(X_1,X_2;Y_1|Q), \tag{12c}$$

taken over the union of all $P_{Q,X_1,X_2} = P_Q P_{X_1|Q} P_{X_2|Q}$, is achievable. We see that the single rate bounds in (12) match the upper bounds in (11). We next intend to show that when the condition in (10) holds, the sum-rate bound in (12c) is redundant. By summing (12a) and (12b)

$$\begin{aligned} R_1 + R_2 &\le I(X_1; Y_1|X_2, Q) + I(X_2; Y_2|Q) \\ &\overset{a)}{\le} I(X_1; Y_1|X_2, Q) + I(X_2; Y_2, X_1|Q) \\ &\overset{b)}{=} I(X_1; Y_1|X_2, Q) + I(X_2; Y_2|X_1, Q) \\ &\overset{c)}{\le} I(X_1; Y_1|X_2, Q) + I(X_2; Y_1|Q) \\ &= I(X_1, X_2; Y_1|Q) = \text{eq.(12c)}, \end{aligned}$$

where in a) we loosened the achievable sum-rate by adding X_1 as "side information" to receiver 2, in b) we used the independence of the inputs, and in c) the condition in (10). Therefore, the sum-rate bound in (12c) can be dropped without affecting the achievable rate region. This shows that the outer bound in (11) is achievable thereby proving the claimed capacity result. ■

Remark 2. *For the classical IC, the very strong interference regime is defined as*

$$I(X_1; Y_1|X_2) \le I(X_1; Y_2),$$
$$I(X_2; Y_2|X_1) \le I(X_2; Y_1),$$

for all product input distributions; under these pair of conditions capacity can be shown. For the IC-OR, the very strong interference constraint at receiver 2 (oblivious receiver) is not need in order to show capacity. Therefore, the very strong interference condition for the IC-OR is less stringent than that for the classical IC. We believe this is so because the oblivious receiver (receiver 2) cannot decode the message of user 1 as per the modeling assumption. Indeed, we feel that the "lack of codebook knowledge" as originally proposed in [3] actually models the inability of a receiver to jointly decode its message along with unintended ones, as the mapping between the messages and codewords is not known. However, we believe that the modeling assumption does not preclude the ability to estimate the codewords of the user whose codebook index is not known.

C. Capacity to within a Constant Gap for the ISD-IC-OR

We now show that $\mathcal{R}_{\text{in}}$ in Proposition 3 lies to within a constant gap of the outer bound $\mathcal{R}_{\text{out}}$ in Theorem 2 for the general ISD-IC-OR. We have

Theorem 5. *For the ISD-IC-OR, if* $(R_1, R_2) \in \mathcal{R}_{\rm out}$ *then* $([R_1 - I(X_2; T_2|U_2, Q)]^+, R_2) \in \mathcal{R}_{\rm in}$.

Proof of Theorem 5: The proof is as in [11]. First, wdefine a new outer bound region $\bar{\mathcal{R}}_{\rm out}$ by replacing X_2 with U_2 in all *positive* entropy terms of region $\mathcal{R}_{\rm out}$, which is permitted as $H(Y_2|X_2) \leq H(Y_2|U_2)$ by the data processing inequality. We conclude that $\mathcal{R}_{\rm out} \subseteq \bar{\mathcal{R}}_{\rm out}$. We next compare $\bar{\mathcal{R}}_{\rm out}$ and $\mathcal{R}_{\rm in}$ term by term (we only need to compare the mutual informations invoking Y_1 as those involving Y_2 are the same in both bounds, see Remark 1, thus implying a zero gap for rate R_2): the difference is that $\bar{\mathcal{R}}_{\rm out}$ has $-H(Y_1|X_1, X_2)$ where $\mathcal{R}_{\rm in}$ has $-H(T_2|U_2, Q)$; thus the gap is

$$-H(Y_1|X_1, X_2) + H(T_2|U_2, Q) = -H(T_2|X_2) + H(T_2|U_2, Q) = I(X_2; T_2|U_2, Q).$$

This concludes the proof. ■

We will give an example of a constant gap characterization in Section IV-E after having discussed in Section IV-D a special class of ISD-IC-OR for which the gap to capacity is zero.

D. Exact Capacity for the Injective Fully Deterministic IC-OR

We now specialize Theorem 5 to the class of injective *fully deterministic* ICs [16]. For this class of channels the mappings T_1 and T_2 in (5) are deterministic functions of X_1 and X_2, respectively. We have

Corollary 6. *For the injective fully deterministic IC-OR the outer bound in Theorem 2 is tight.*

Proof of Corollary 6: The injective fully deterministic IC-OR has $T_2 = U_2$ and therefore $I(X_2; T_2|U_2, Q) = 0$ in Theorem 5. ■

As an application of Corollary 6 we consider next the Linear Deterministic Approximation (LDA) of the Gaussian IC-OR at high SNR, whose classical counterpart (where all codebooks are known) was first proposed in [12]. The LDA-IC-OR has input/output relationship

$$Y_1 = \mathbf{S}^{q-n_{11}} X_1 + \mathbf{S}^{q-n_{12}} X_2, \quad T_2 = \mathbf{S}^{q-n_{12}} X_2, \tag{13a}$$

$$Y_2 = \mathbf{S}^{q-n_{21}} X_1 + \mathbf{S}^{q-n_{22}} X_2, \quad T_1 = \mathbf{S}^{q-n_{21}} X_1, \tag{13b}$$

where inputs and outputs are binary-valued vectors of length q, $\mathbf{S}$ is the $q \times q$ shift matrix [12], $(n_{11}, n_{12}, n_{21}, n_{22})$ are non-negative integers and $q := \max\{n_{11}, n_{12}, n_{21}, n_{22}\}$. Summations and multiplications are bit-wise over the binary field.

For simplicity, we next evaluate the *symmetric* sum-capacity of the LDA-IC-OR. The symmetric LDA-IC-OR has parameters $n_{11} = n_{22} = n_{\mathrm{S}}$ and $n_{12} = n_{21} = n_{\mathrm{I}} := n_{\mathrm{S}}\ \alpha$ for some non-negative α. The maximum symmetric rate, or symmetric sum-capacity normalized by the sum-capacity of an interference-free channel, is defined as

$$d(\alpha) := \frac{\max\{R_1 + R_2\}}{2\ n_{\mathrm{S}}} \tag{14}$$

where the maximization is over all achievable rate pairs (R_1, R_2) satisfying Theorem 2, which is the capacity region by Corollary 6. Since we may provide the oblivious receiver in the LDA-IC-OR with the additional codebook index so as to obtain the classical LDA-IC with full codebook knowledge, we immediately have

$$d(\alpha) \leq d^{(\mathrm{W})}(\alpha) = \min\left(1, \max\left(\frac{\alpha}{2}, 1 - \frac{\alpha}{2}\right), \max\left(\alpha, 1 - \alpha\right)\right), \tag{15}$$

where $d^{(\mathrm{W})}(\alpha)$, the so-called W-curve [17], is the maximum symmetric rate of the classical LDA-IC. In [13] it was shown that i.i.d. $\mathsf{Bernoulli}(1/2)$ input bits in the Han and Kobayashi region yield $d^{(\mathrm{W})}(\alpha)$.

Although Theorem 2 gives the exact capacity region of the LDA-IC-OR, it is not immediately clear which input distribution achieves the maximum symmetric rate. Instead of analytically deriving the sum-capacity, we proceeded to numerically evaluate Theorem 2 for $|Q| = 1$, which is not necessarily optimal. We observed the surprising result that even with $|Q| = 1$, i.e., without time sharing, the normalized sum-capacity of the LDA-IC-OR equals $d^{(\mathrm{W})}(\alpha)$, see Fig. 2 and Table I. This implies that partial codebook knowledge at one receiver does not impact the sum-rate of the symmetric LDA-IC-OR. This is quite unexpected, especially in the strong interference regime ($\alpha \geq 1$) where the optimal strategy for the classical LDA-IC is to jointly decode the interfering message along with the intended message—a strategy that seems to be precluded by the lack of codebook knowledge at one receiver. *This might suggest a more general principle: there is no loss of optimality in lack of codebook knowledge as long as the oblivious receiver can remove the interfering codeword, regardless of whether or not it can decode the message carried by this codeword.*

Another interesting observation is that i.i.d. $\mathsf{Bernoulli}(1/2)$ input bits may no longer be optimal (though we not not show their strict sub-optimality). In Table I we report, for some values of α and n_{S}, the input distributions to be used in $\mathcal{R}_{\mathrm{out}}$ in Theorem 2. We notice that, at least

TABLE I

LDA-IC-OR: EXAMPLES OF SUM-RATE OPTIMAL INPUT DISTRIBUTIONS FOR THE CAPACITY REGION IN THEOREM 2.

α, $n_{\rm S}$	Probability mass function with $|Q|=1$
$\frac{1}{2}$, 2	$P_{X_1}=[0.5,0,0.5,0]$
	$P_{X_2}=[0,0.5,0,0.5]$
$\frac{2}{3}$, 3	$P_{X_1}=[0,0,0.25,0.25,0,0,0.25,0.25]$
	$P_{X_2}=[0,0,0.25,0.25,0,0,0.25,0.25]$
1, 2	$P_{X_1}=[0,0,0.5,0.5]$
	$P_{X_2}=[0,0.5,0,0.5]$
$\frac{4}{3}$, 4	$P_{X_1}=[0,0,0,0,0,0.25,0,0.25,0,0,0,0,0,0.25,0,0.25]$
	$P_{X_2}=[0,0,0,0.25,0,0.25,0,0,0,0,0,0,0,0.25,0,0.25]$
2, 2	$P_{X_1}=[0,0.5,0,0.5]$
	$P_{X_2}=[0,0.5,0,0.5]$

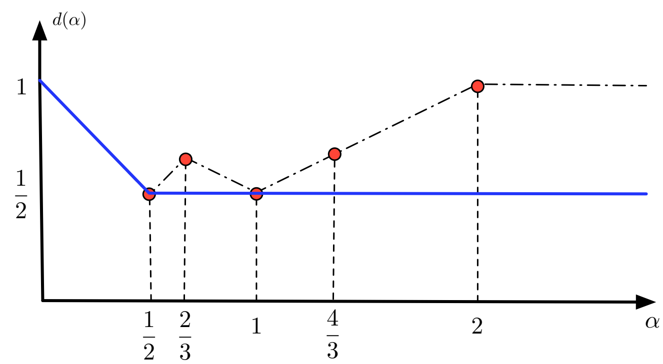


Fig. 2. The normalized sum-capacity, or maximum symmetric rate, for the classical LDA-IC (dash-dotted black line). Normalized sum-rates achieved by the input distributions in Table I (red dots) for the LDA-IC-OR. The normalized sum-rate achieved by i.i.d. $\mathsf{Bernoulli}(1/2)$ inputs and $|Q|=1$ (solid blue line) in the capacity region in Theorem 2 for the LDA-IC-OR.

when evaluating the region in Theorem 2 for $|Q|=1$ only, that the region exhausting inputs are now *correlated*. For example, Table I shows that, for $\alpha=4/3$ and $n_{\rm S}=4$, the inputs X_1 and X_2 are binary vectors of length $n_{\rm S}=\log(16)=4$ bits; out of the 16 different possible bit sequences, only 4 are actually used at each transmitter with strictly positive probability to achieve $d^{\rm (W)}(4/3)=4/6$. By using i.i.d. $\mathsf{Bernoulli}(1/2)$ input bits in Theorem 2 for $|Q|=1$ we would obtain a normalized sum-rate of $1/2=3/6$, the same as achieved by time division [13].

E. The Gaussian Noise IC-OR

We now consider the practically relevant real-valued single-antenna power-constrained Gaussian noise channel, whose input/output relationship is

$$Y_1 = h_{11}X_1 + h_{12}X_2 + Z_1 = h_{11}X_1 + T_2, \quad T_2 = h_{12}X_2 + Z_1, \tag{16a}$$

$$Y_2 = h_{21}X_1 + h_{22}X_2 + Z_2 = h_{22}X_2 + T_1, \quad T_1 = h_{21}X_1 + Z_2, \tag{16b}$$

where h_{ij} are the real-valued channel coefficients for $(i, j) \in [1 : 2]^2$ assumed constant and known to all nodes, the input $X_i \in \mathbb{R}$ is subject to per block power constraints $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \leq 1,\ i \in [1 : 2]$, and the noise $Z_i,\ i \in [1 : 2]$, is a unit-variance zero-mean Gaussian r.v..

By specializing the result of Theorem 5 to the G-IC-OR we may show the following:

Corollary 7. *For the G-IC-OR the gap is at most* $1/2$ *bit (per channel use per user).*

Proof of Corollary 7: For the G-IC-OR $T_2 = h_{12}X_2 + Z_1$, and thus we set U_2 in Theorem 2 to $U_2 = h_{12}X_2 + Z_1^*$, where $Z_1 \sim Z_1^*$ and mutually independent. We thus have

$$\begin{aligned} I(X_2; T_2|U_2, Q) &= h(T_2|U_2, Q) - h(Z_2) \\ &\leq h(T_2 - U_2) - h(Z_1) \\ &= h(Z_1 - Z_1^*) - h(Z_1) = \frac{1}{2}\log(2), \end{aligned}$$

as claimed. ■

In the classical G-IC with full codebook knowledge, Gaussian inputs exhaust known outer bounds, which are achievable to within $1/2$ bit (per channel use per user) [17]. From the rate expression in Theorem 2 it is not clear whether Gaussian inputs are optimal for $\mathcal{R}_{\text{out}}$. The following discussion shows that in general the answer is in the negative. For simplicity we focus on the achievable generalized Degrees of Freedom (gDoF) for the symmetric G-IC-OR. The symmetric G-IC-OR has $|h_{11}|^2 = |h_{22}|^2 = \mathsf{SNR}$ and $|h_{12}|^2 = |h_{21}|^2 = \mathsf{INR}$, with $\mathsf{INR} = \mathsf{SNR}^\alpha$ for some non-negative α. The sum-gDoF is defined as

$$d(\alpha) := \lim_{\mathsf{SNR}\to+\infty} \frac{\max\{R_1 + R_2\}}{2 \cdot \frac{1}{2}\log(1 + \mathsf{SNR})}, \tag{17}$$

where the maximization is over all possible achievable rate pairs. By using the classical G-IC as a trivial upper bound, we have $d(\alpha) \leq d^{(\mathrm{W})}(\alpha)$ where $d^{(\mathrm{W})}(\alpha)$ is given in (15).

By evaluating Theorem 2 for independent Gaussian inputs and $|Q| = 1$ (which we do *not* claim to be optimal, but which gives us an achievable rate up to $1/2$ bit) we obtain

$$(R_1 + R_2)^{(\mathrm{GG})} = \min\left\{\mathsf{I_g}\left(\mathsf{SNR}\right) + \mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \right.$$
$$\left. \mathsf{I_g}\left(\frac{\mathsf{SNR}}{\mathsf{INR}+1}\right) + \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)\right\},$$
$$\iff \quad d^{(\mathrm{GG})}(\alpha) = \frac{1}{2} + \left[\frac{1}{2} - \alpha\right]^+,$$

the superscript "GG" indicates that both transmitters use a Gaussian input. For future reference, with Time Division (TD) and Gaussian codebooks we can achieve

$$(R_1 + R_2)^{(\mathrm{TD})} = \frac{1}{2}\log\left(1 + 2\ \mathsf{SNR}\right) \quad \iff \quad d^{(\mathrm{TD})}(\alpha) = \frac{1}{2}.$$

We plot the achievable gDoF vs. α in Fig. 2, together with the gDoF of the classical G-IC given by $d^{(\mathrm{W})}(\alpha)$ [17]. We note that Gaussian inputs are indeed optimal for $0 \le \alpha \le 1/2$, i.e., $d^{(\mathrm{GG})}(\alpha) = d^{(\mathrm{W})}(\alpha)$, where interference is treated as noise even for the classical G-IC (which is also achievable by the G-IC-OR). For $\alpha > 1/2$ we have $d^{(\mathrm{GG})}(\alpha) = d^{(\mathrm{TD})}(\alpha)$, that is, Gaussian inputs achieve the same rates as time division. Interestingly, Gaussian inputs are sub-optimal in our achievable region in general as we show next.

Consider $\alpha = 4/3$. With Gaussian inputs we only achieve $d^{(\mathrm{GG})}(4/3) = d^{(\mathrm{TD})}(4/3) = 1/2$. Notice the similarity with the LDA-IC-OR: the input distribution that is optimal for the non-oblivious IC performs as time division for the G-IC-OR. Inspired by the LDA-IC-OR we explore now the possibility of using non-Gaussian inputs. By following [3, Section VI.A], which demonstrated that binary signaling outperforms Gaussian signaling for a fixed finite SNR, we consider a uniform PAM constellation with N points. Fig. 3 shows the achievable normalized sum-rate $\frac{R_1+R_2}{2\cdot\frac{1}{2}\log(1+\mathsf{SNR})}$ as a function of SNR for the case where X_1 (the input of the non-oblivious pair) is a PAM constellation with $N = \left\lfloor \mathsf{SNR}^{1/6} \right\rfloor$ points and X_2 (the input of the oblivious pair) is Gaussian; we refer to the achievable gDoF of this inputs as $d^{(\mathrm{DG})}(\alpha)$. Notice that the number of points in the discrete input is a function of SNR. We also report the achievable normalized sum-rate with time division and Gaussian inputs. Fig. 3 shows that, for sufficiently large SNR, using a discrete input outperforms time division; moreover, for the range of simulated SNR, it seems that the proposed discrete input achieves a gDoF of $d^{(\mathrm{DG})}(\alpha) = \alpha/2 = 4/6$ as for the classical G-IC with full codebook knowledge. In the sections that follow we analytically show

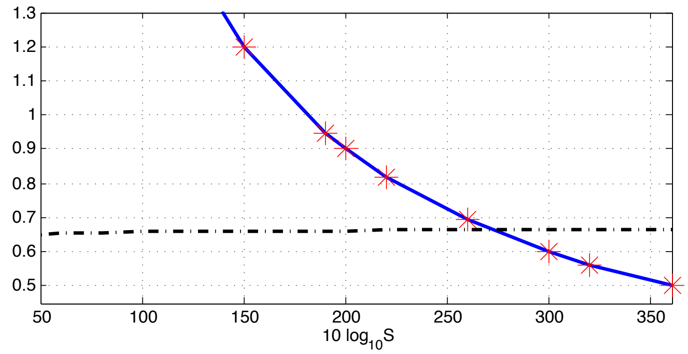


Fig. 3. Achievable normalized sum-rate for the symmetric G-IC-OR with $\alpha = 4/3$ vs SNR in dB. Legend: time division in solid blue line; Gaussian inputs at both transmitters in red stars; X_1 is a uniform PAM with $N = \lfloor \mathsf{SNR}^{\frac{1}{6}} \rfloor$ points and X_2 is Gaussian in dash-dotted black line.

that using discrete input (or mixed) at the non-oblivious transmitter indeed achieves the full gDoF and capacity to within a constant gap.

V. Discrete Inputs: main tool

In this section we review the lower bound of [7] on the mutual information achievable by a PAM input on a point-to-point power-constrained Gaussian noise channel that will serve as the main tool in evaluating our inner bound for the G-IC-OR in Proposition 3. The bound is as follows.

Theorem 8. *Let* $X_D \sim \mathsf{PAM}(N)$ *and let* $Z_G \sim \mathcal{N}(0,1)$ *and* SNR *be a non-negative constant. Then*

$$\left[\mathsf{I_g}\left(\min\left(N^2-1, \mathsf{SNR}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+ =: \mathsf{I_d}\left(N, \mathsf{SNR}\right) \tag{18}$$

$$\leq I(X_D; \sqrt{\mathsf{SNR}}\, X_D + Z_G) \leq \mathsf{I_g}\left(\min\left(N^2-1, \mathsf{SNR}\right)\right). \tag{19}$$

Proof of Theorem 8: The upper bound in (19) follows from the well known facts that "Gaussian maximizes the differential entropy for a given second moment constraint" and that "a uniform input maximizes the entropy of a discrete random variable" [14]. Let now $x_{\min} :=$

$\min(N^2-1,\mathsf{SNR})$ and $x_{\max} := \max(N^2-1,\mathsf{SNR})$. We have

$$
\begin{aligned}
& I(X_D; \sqrt{\mathsf{SNR}}X_D + Z_G) \\
& \overset{\text{from [7, Part b]}}{\geq} \frac{1}{2}\log\left(1+(N^2-1)\right) - \frac{1}{2}\log\left(1+\frac{N^2-1}{1+\mathsf{SNR}}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{6}\right) \\
& = \mathsf{I_g}\left(x_{\min}\right) + \mathsf{I_g}\left(x_{\max}\right) - \mathsf{I_g}\left(x_{\min}+x_{\max}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{6}\right) \\
& = \mathsf{I_g}\left(x_{\min}\right) - \mathsf{I_g}\left(\frac{x_{\min}}{1+x_{\max}}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{6}\right) \\
& \geq \mathsf{I_g}\left(x_{\min}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right),
\end{aligned}
$$

since $\frac{x_{\min}}{1+x_{\max}} \in [0,1]$. This, combined with non-negativity of mutual information, gives the lower bound in (18). ■

Remark 3. *The upper and lower bounds in Theorem 8 are to within* $\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)$ *bits of one another. We shall refer to the quantity* $\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)$ *as the "shaping loss" due to the use of a one-dimensional lattice constellation on the power-constrained point-to-point Gaussian channel. Note that what is known as "shaping gain" of a one-dimensional lattice constellation in the literature is* $\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{6}\right)$ *[18]; what we call here "shaping loss" has an extra* $\frac{1}{2}\log(2)$ *due to the term* $\mathsf{I_g}\left(\frac{x_{\min}}{1+x_{\max}}\right)$*; we refer to the sum of these two contributions as "shaping loss" because it is purely due to the one-dimensional lattice ("shaping" part) and it causes a reduction in rate compared to the upper bound ("loss" part).*

If we could choose $N^2-1=\mathsf{SNR} \iff N=\sqrt{1+\mathsf{SNR}}$ *then we could claim that a PAM input is optimal (i.e., achieves the capacity of the point-to-point power-constrained Gaussian noise channel) to within* $\mathsf{gap} \leq \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)$ *bits per channel use, where the gap would be completely due to the shaping loss.*

Unfortunately, N *is constrained to be an integer. If for* N *we choose the closest integer to* $\sqrt{1+\mathsf{SNR}}$*, that is,* $N = \left\lfloor\sqrt{1+\mathsf{SNR}}\right\rfloor =: \mathsf{N_d}(\mathsf{SNR})$*, then we incur a further 1 bit "integer penalty", by which we mean that the difference between the point-to-point Gaussian channel capacity and the lower bound on the achievable rate with a PAM in* (18) *is upper bounded as*

$$
\begin{aligned}
\mathsf{gap} & \leq \mathsf{I_g}\left(\mathsf{SNR}\right) - \mathsf{I_d}\left(\mathsf{N_d}(\mathsf{SNR}),\mathsf{SNR}\right) \\
& \leq \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log^{+}\left(\frac{1+\mathsf{SNR}}{\left\lfloor\sqrt{1+\mathsf{SNR}}\right\rfloor^2}\right)
\end{aligned}
$$

$$\leq \underbrace{\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)}_{\textit{shaping loss}} + \underbrace{\frac{1}{2}\log(4)}_{\textit{integer penalty}} = \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right), \tag{20}$$

where the largest integer penalty is attained for $1+\mathsf{SNR} = 2^2-\epsilon,\ 0<\epsilon\ll 1$*, for which* $\lfloor\sqrt{1+\mathsf{SNR}}\rfloor^2 = (2-1)^2 = 1$*. An interesting question is by how much the "shaping loss" and the "integer penalty" can be reduced by using multidimensional "good" lattices [19].*

VI. Achievable Regions for the G-IC-OR

We now analyze the G-IC-OR by using Theorem 8 (i.e., bounds on the mutual information achievable by a PAM input on a point-to-point power-constrained Gaussian noise channel) and the insight on the nature of the gap due to a PAM input from Remark 3. We first present a scheme (an achievable rate region evaluated using a mixed input) that will prove to be useful in strong and very strong interference, and then present a more involved scheme that will be useful in the somewhat trickier weak and moderate interference regimes. Although the second scheme includes the first as a special case, we start with a simpler scheme to highlight the important steps of the derivation without getting caught up in excessive technical details.

A. Achievable Scheme I

We first derive an achievable rate region from Proposition 3 with inputs

$$\text{Scheme I:}\quad X_{1D} \sim \mathsf{PAM}\left(N\right),\ N\in\mathbb{N},\ \text{independent of} \tag{21a}$$

$$X_{2G} \sim \mathcal{N}(0,1), \tag{21b}$$

$$X_1 = X_{1D},\ X_2 = X_{2G}, \tag{21c}$$

$$U_2 = X_2,\ Q = \emptyset. \tag{21d}$$

which we will show in the next sections to be gDoF optimal and to within a constant gap of the symmetric capacity of the classical G-IC in the strong and very strong interference regimes. Such results may not be shown by using i.i.d. Gaussian inputs in the same achievable scheme in Proposition 3. The achievable region is derived for a general G-IC-OR and later on specialized to the symmetric case.

Theorem 9. *For the G-IC-OR the following rate region is achievable by the input in* (21)

$$R_1 \leq \mathsf{I_d}\left(N, |h_{11}|^2\right), \tag{22a}$$

$$R_2 \leq \mathsf{I_d}\left(N, \frac{|h_{21}|^2}{1+|h_{22}|^2}\right) + \mathsf{I_g}\left(|h_{22}|^2\right) - \mathsf{I_g}\left(\min\left(N^2-1, |h_{21}|^2\right)\right), \tag{22b}$$

$$R_1 + R_2 \leq \mathsf{I_d}\left(N, \frac{|h_{11}|^2}{1+|h_{12}|^2}\right) + \mathsf{I_g}\left(|h_{12}|^2\right). \tag{22c}$$

Proof of Theorem 9: We proceed to evaluate the rate region in Proposition 3 with the inputs in (21), that is, the achievable region in (12) with $|Q| = 1$.

The rate of the user 1 is bounded by $R_1 \leq I(X_1; Y_1|X_2) = I(X_{1D}; h_{11}X_{1D} + Z_1)$, where $I(X_{1D}; h_{11}X_{1D} + Z_1)$ can be further lower bounded by using (18) from Theorem 8 with $\mathsf{SNR} = |h_{11}|^2$; by doing so we obtain the bound in (22a).

The rate of the user 2 is bounded by

$$\begin{aligned}
R_2 &\leq I(X_2; Y_2) \\
&= h(h_{21}X_{1D} + \underbrace{h_{22}X_{2G} + Z_2}_{\sim\mathcal{N}(0,1+|h_{22}|^2)}) - h(h_{21}X_{1D} + Z_2) \\
&= \underbrace{\left(h\left(\frac{h_{21}}{\sqrt{1+|h_{22}|^2}}X_{1D} + Z_2\right) - h(Z_2)\right)}_{\geq \mathsf{I_d}\left(N, \frac{|h_{21}|^2}{1+|h_{22}|^2}\right) \text{ from (18)}} + \frac{1}{2}\log(1+|h_{22}|^2) \\
&\quad - \underbrace{\left(h\left(h_{21}X_{1D} + Z_2\right) - h(Z_2)\right)}_{\leq \mathsf{I_g}(\min(N^2-1, |h_{21}|^2)) \text{ from (19)}},
\end{aligned}$$

from which we conclude that the achievable rate for user 2 is lower bounded as in (22b).

The sum-rate is bounded by $R_1 + R_2 \leq I(X_1, X_2; Y_1) = I(X_1; Y_1) + I(X_2; Y_1|X_1)$, where where $I(X_1; Y_1)$ can be lower bounded by means of Theorem 8 with $\mathsf{SNR} = \frac{|h_{11}|^2}{1+|h_{12}|^2}$ and where $I(X_2; Y_1|X_1) = I(X_{2G}; h_{12}X_{2G} + Z_1) = \mathsf{I_g}(|h_{12}|^2)$; by combining the two terms we obtain the bound in (22c). ■

B. Achievable Scheme II

The input in (21) might not be optimal in general and may be generalized as follows. Consider the rate region in Proposition 3 with inputs

$$\text{Scheme II:} \quad X_{1D}, X_{1G}, X_{2Gc}, X_{2Gp} \text{ independent and distributed as} \tag{23a}$$

$$X_{1D} \sim \mathsf{PAM}\,(N)\,, \;\; N \in \mathbb{N}, \tag{23b}$$

$$\text{all the others are } \mathcal{N}(0,1), \tag{23c}$$

$$X_1 = \sqrt{1-\delta_1} X_{1D} + \sqrt{\delta_1} X_{1G}, \;\; \delta_1 \in [0,1], \tag{23d}$$

$$X_2 = \sqrt{1-\delta_2} X_{2Gc} + \sqrt{\delta_2} X_{2Gp}, \;\; \delta_2 \in [0,1]. \tag{23e}$$

$$U_2 = X_{2Gc}, \;\; Q = \emptyset. \tag{23f}$$

In Scheme II, X_{2Gc} encodes a “common” message, and X_{2Gp} and X_{1G} encode the “private” messages as in the classical Han-Kobayashi scheme [6]. We shall also interpret X_{1D} as encoding a “common” message even if X_{1D} cannot be decoded at receiver 2 (the oblivious receiver) as receiver 2 lacks knowledge of the codebook(s) used by transmitter 1. The main message of the paper is in fact that, even with lack of codebook knowledge, if there would-be-common message is from a discrete alphabet then its effect on the rate region—up to a constant gap—is as if the message could indeed be jointly decoded. We believe this is because lack of codebook knowledge may be translated as lack of knowledge of the mapping of the codewords to the messages, but does not preclude a receiver's ability to *estimate* the interfering codeword (rather than messages). “Estimating” and subtracting off the interfering codeword is as effective in terms of rates as decoding the actual interfering message encoded in the given codeword, as this message is not desired anyhow.

In the next sections we will show that Proposition 3 with the inputs in (23) is gDoF optimal and is to within a constant gap of a capacity outer bound for the classical G-IC in the weak and moderate interference regimes. Also note that with $\delta_1 = \delta_2 = 0$ Scheme II in (23) reduces to Scheme I in (21).

The achievable region is derived for a general G-IC-OR and later on specialized to the symmetric case. The rate region achievable by Scheme II is

Theorem 10. *For the G-IC-OR the following rate region is achievable with inputs as in* (23)

$$R_1 \le \mathsf{I_d}\left(N, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2\delta_2}\right) + \mathsf{I_g}\left(\frac{|h_{11}|^2\delta_1}{1+|h_{12}|^2\delta_2}\right), \tag{24a}$$

$$R_2 \le \mathsf{I_d}\left(N, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2}\right) + \mathsf{I_g}\left(\frac{|h_{22}|^2}{1+|h_{21}|^2\delta_1}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right), \tag{24b}$$

$$R_1+R_2 \le \mathsf{I_d}\left(N, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2}\right) + \mathsf{I_g}\left(|h_{11}|^2\delta_1+|h_{12}|^2\right) - \mathsf{I_g}\left(|h_{12}|^2\delta_2\right) + \mathsf{I_d}\left(N, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2\delta_2}\right) + \mathsf{I_g}\left(\frac{|h_{22}|^2\delta_2}{1+|h_{21}|^2\delta_1}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right). \tag{24c}$$

Proof of Theorem 10: The proof can be found in Appendix A and follows similarly to the proof of Theorem 9. ■

VII. High SNR performance

We now analyze the performance of the schemes in Theorems 9 and 10 for the symmetric G-IC-OR at high-SNR by using the gDoF region as performance metric. The gDoF region is formally defined as follows. For an achievable pair (R_1, R_2), let

$$\mathcal{D}(\alpha) := \left\{ (d_1, d_2) \in \mathbb{R}_+^2 : d_i := \lim_{\substack{\mathsf{INR} = \mathsf{SNR}^\alpha, \\ \mathsf{SNR} \to \infty}} \frac{R_i}{\frac{1}{2}\log(1+\mathsf{SNR})}, i \in [1:2],\ (R_1, R_2) \text{ is achievable} \right\}. \tag{25}$$

Let $\mathcal{D}^{\text{G-IC}}(\alpha)$ and $\mathcal{D}^{\text{G-IC-OR}}(\alpha)$ be the gDoF region of the classical G-IC and of the G-IC-OR, respectively. The key novelty of our approach is to take $N = \mathsf{N_d}(\mathsf{SNR}^\beta)$ for some $\beta \ge 0$. We first present two different achievable gDoF regions based on Theorems 9 and 10, which we will compare to $\mathcal{D}^{\text{G-IC}}(\alpha)$ given by [17]

$$\mathcal{D}^{\text{G-IC}}(\alpha) \ : \ d_1 \le 1, \tag{26a}$$

$$d_2 \le 1, \tag{26b}$$

$$d_1 + d_2 \le \max(\alpha, 2-\alpha), \tag{26c}$$

$$d_1 + d_2 \le \max(2\alpha, 2-2\alpha), \tag{26d}$$

$$2d_1 + d_2 \le 2, \ \text{only for } \alpha \in [1/2, 1], \tag{26e}$$

$$d_1 + 2d_2 \le 2, \ \text{only for } \alpha \in [1/2, 1]. \tag{26f}$$

Corollary 11 (gDoF region from achievable Scheme I)**.** *Let*

$$\mathcal{D}^{\text{I}}(\alpha,\beta) \ : \ d_1 \le \min(\beta, 1), \tag{27a}$$

$$d_2 \le \min(\beta, [\alpha-1]^+) + 1 - \min(\beta, \alpha), \tag{27b}$$

$$d_1 + d_2 \le \min(\beta, [1-\alpha]^+) + \alpha. \tag{27c}$$

for any $\beta \ge 0$. By Theorem 9, the gDoF region $\mathcal{D}^{\text{I}}(\alpha,\beta)$ is achievable.

Proof of Corollary 11: We prove the bound in (27b) only as the other bounds follow similarly. With $\mathsf{INR} = \mathsf{SNR}^\alpha$ and $N = \mathsf{N_d}(\mathsf{SNR}^\beta)$ we have

$$\lim_{\mathsf{SNR}\to\infty} \frac{\log(N^2)}{\log(1+\mathsf{SNR})} = \beta,$$
$$\lim_{\mathsf{SNR}\to\infty} \frac{\log(1+\mathsf{INR})}{\log(1+\mathsf{SNR})} = \alpha.$$

Therefore d_2 can be bounded as

$$\begin{aligned} d_2 &= \lim_{\mathsf{SNR}\to\infty} \frac{\text{left hand side of eq.(22b)}}{\frac{1}{2}\log(1+\mathsf{SNR})} \\ &= \min(\beta, [\alpha-1]^+) + 1 - \min(\beta,\alpha), \end{aligned}$$

thus proving (27b). ■

Next, by using Theorem 10 with the power split as in [17] we show yet another achievable gDoF region.

Corollary 12 (gDoF region from achievable Scheme II)**.** *Let*

$$\mathcal{D}^{\rm II}(\alpha,\beta) \ : \ d_1 \leq \min(\beta, 1+\alpha-\max(1,\alpha)) + [1-\alpha]^+, \tag{28a}$$

$$d_2 \leq \min(\beta, [\alpha-1]^+) + 1 - \min(\beta,\alpha), \tag{28b}$$

$$d_1 + d_2 \leq \min(\beta, [1+\alpha-\max(1,2\alpha)]^+) + \max(\alpha, 1-\alpha) + $$
$$+ \min(\beta, [2\alpha - \max(1,\alpha)]^+) + [1-\alpha]^+ - \min(\beta,\alpha). \tag{28c}$$

for any $\beta \geq 0$. By Theorem 10, the gDoF region $\mathcal{D}^{\rm II}(\alpha,\beta)$ is achievable.

Proof of Corollary 12: Let $\mathsf{INR} = \mathsf{SNR}^{\alpha}$, $N = \mathsf{N_d}(\mathsf{SNR}^{\beta})$, and $\delta_1 = \delta_2 = \frac{1}{1+\mathsf{INR}}$ in Theorem 10 (see the region in (41) in Appendix A) and take limits similarly to the proof of Corollary 11. ■

We are now ready to prove the main result of this section:

Theorem 13. *For the G-IC-OR there is no loss in gDoF compared to the classical G-IC, i.e.,*

$$\mathcal{D}^{\textit{G-IC}}(\alpha) = \mathcal{D}^{\textit{G-IC-OR}}(\alpha).$$

Proof of Theorem 13: We consider several regimes:

a) Very strong interference regime $\alpha \geq 2$: In this regime the gDoG region outer bound $\mathcal{D}^{\text{G-IC}}(\alpha)$ is characterized by (26a) and (26b). For achievability we consider Corollary 11 with $\beta = 1$, that is,

$$\mathcal{D}^{\rm I}(\alpha,1) \ : \ d_1 \leq \min(1,1) = 1,$$

$$d_2 \leq \min(1, [\alpha-1]^+) + 1 - \min(1,\alpha) = 1,$$

$$d_1 + d_2 \leq \min(1, [1-\alpha]^+) + \alpha = \alpha \ \text{(redundant because } \alpha \geq 2).$$

Since the sum-gDoF is redundant, we get that

$$\mathcal{D}^{\rm I}(\alpha, \beta=1) = \{d_i \in [0,1], \ i \in [1:2]\} = \mathcal{D}^{\text{G-IC-OR}}(\alpha) = \mathcal{D}^{\text{G-IC}}(\alpha).$$

Fig. 4(a) illustrates the region $\mathcal{D}^{\rm I}(\alpha, \beta=1)$.

b) Strong interference regime $1 \leq \alpha < 2$: In this regime the gDoG region outer bound $\mathcal{D}^{\text{G-IC}}(\alpha)$ is characterized by (26a)-(26c) and has two dominant corner points: $(d_1, d_2) = (1, \alpha-1)$ and $(d_1, d_2) = (\alpha - 1, 1)$. For achievability we consider the following achievable gDoF regions

$$\begin{aligned} \mathcal{D}^{\text{I}}(\alpha, 1) \ : \ & d_1 \leq 1, \\ & d_2 \leq \alpha - 1, \\ & d_1 + d_2 \leq \alpha \text{ (redundant)}. \end{aligned}$$

and

$$\begin{aligned} \mathcal{D}^{\text{I}}(\alpha, \alpha-1) \ : \ & d_1 \leq \alpha - 1, \\ & d_2 \leq 1, \\ & d_1 + d_2 \leq \alpha \text{ (redundant)}, \end{aligned}$$

Fig. 4(b) illustrates that

$$\mathsf{co}\left(\mathcal{D}^{\text{I}}(\alpha, 1) \cup \mathcal{D}^{\text{I}}(\alpha, \alpha-1)\right) = \mathcal{D}^{\text{G-IC}}(\alpha) = \mathcal{D}^{\text{G-IC-OR}}(\alpha).$$

c) Moderately weak interference regime $\frac{1}{2} < \alpha < 1$: In this regime the gDoG region outer bound $\mathcal{D}^{\text{G-IC}}(\alpha)$ is characterized by all the constraints in (26) and has four corner points: $(d_1, d_2) = (1, 0)$, $(d_1, d_2) = (0, 1)$, and $(d_1, d_2) = (\min(4\alpha - 2, \alpha), 2 - 2\alpha)$ and $(d_1, d_2) = (2 - 2\alpha, \min(4\alpha - 2, \alpha))$. The gDoF pair $(d_1, d_2) = (1, 0)$ is trivially achievable by silencing user 2, and similarly $(d_1, d_2) = (0, 1)$ by silencing user 1. For achievability of the remaining two corner points, we consider the following achievable gDoF regions

$$\begin{aligned} \mathcal{D}^{\text{II}}(\alpha, 2\alpha-1) \ : \ & d_1 \leq \min(2\alpha - 1, 1 + \alpha - 1) + 1 - \alpha = \alpha, \\ & d_2 \leq \min(2\alpha - 1, 0) + 1 - \min(2\alpha - 1, \alpha) = 2 - 2\alpha, \\ & d_1 + d_2 \leq \min(2\alpha - 1, [1 + \alpha - \max(1, 2\alpha)]^+) + \max(\alpha, 1 - \alpha) + \\ & \quad + \min(2\alpha - 1, [2\alpha - 1]^+) + 1 - \alpha - \min(2\alpha - 1, \alpha) \\ & \quad = \min(2\alpha, 2 - \alpha), \quad \text{(redundant for } \alpha \in [2/3, 1]). \end{aligned}$$

and

$$\begin{aligned}\mathcal{D}^{\mathrm{II}}(\alpha,1-\alpha)\ :\ d_1 &\le \min(1-\alpha,1+\alpha-1)+1-\alpha = 2-2\alpha,\\ d_2 &\le \min(1-\alpha,0)+1-\min(1-\alpha,\alpha)=\alpha,\\ d_1+d_2 &\le \min(1-\alpha,[1+\alpha-\max(1,2\alpha)]^+)+\max(\alpha,1-\alpha)+\\ &+\min(1-\alpha,[2\alpha-1]^+)+1-\alpha-\min(1-\alpha,\alpha)\\ &=\min(2\alpha,2-\alpha),\quad \text{(redundant for } \alpha\in[2/3,1]).\end{aligned}$$

Fig. 4(c) (for $\alpha \in [2/3,1]$) and Fig. 4(d) (for $\alpha \in [1/2,2/3]$) illustrate that

$$\mathsf{co}\left(\left\{(d_1,d_2)=(1,0)\right\}\cup\left\{(d_1,d_2)=(0,1)\right\}\cup\mathcal{D}^{\mathrm{II}}(\alpha,2\alpha-1)\cup\mathcal{D}^{\mathrm{II}}(\alpha,1-\alpha)\right)=\mathcal{D}^{\text{G-IC}}(\alpha)$$
$$=\mathcal{D}^{\text{G-IC-OR}}(\alpha).$$

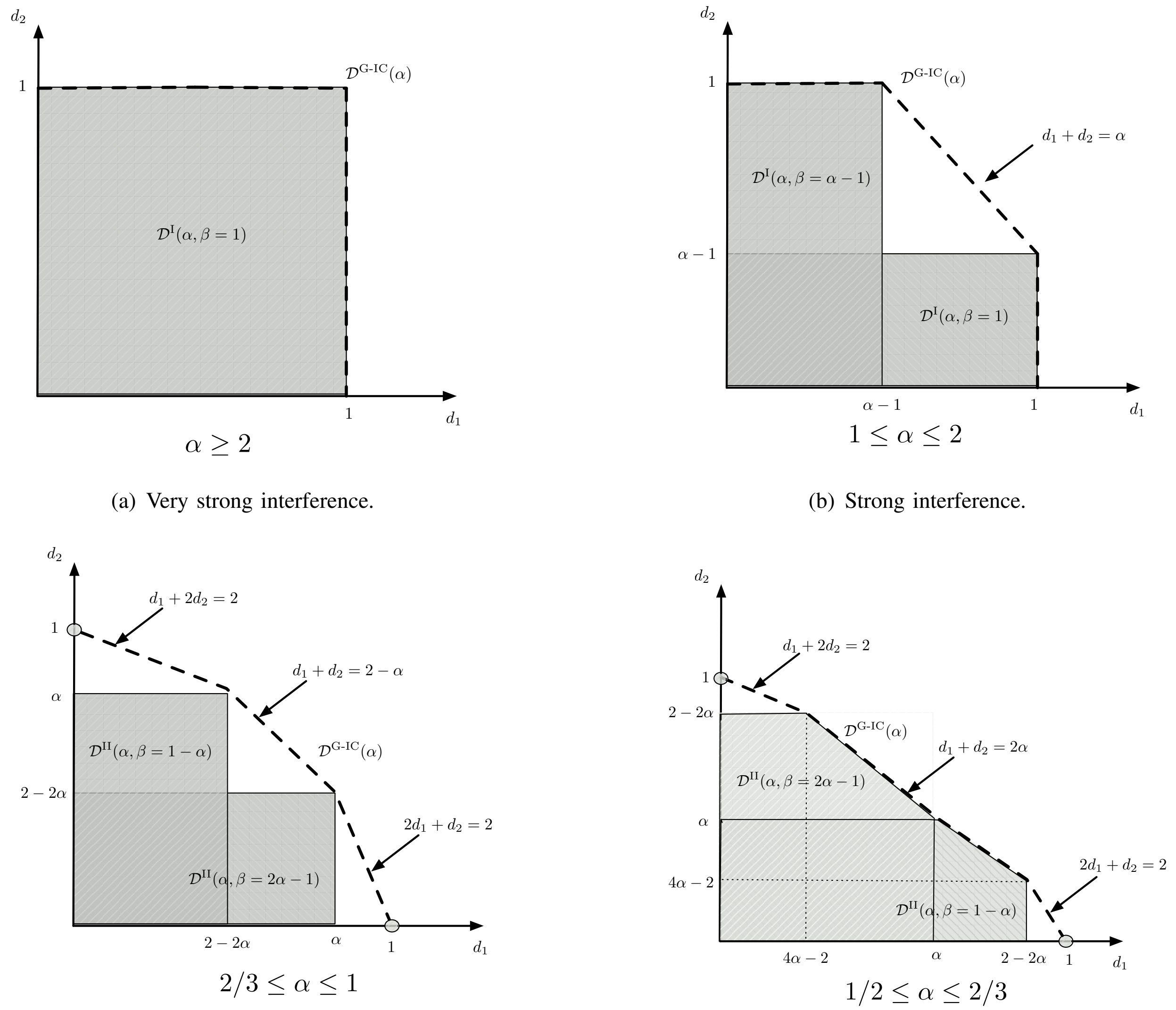
d) Noisy Interference $0 \le \alpha \le \frac{1}{2}$*:* In this regime one may achieve the whole optimal G-IC gDoF region by using Gaussian inputs, treating interference as noise, and power control. Since this strategy is feasible for the G-IC-OR, the G-IC gDoF region is achievable also for the G-IC-OR.

This concludes our proof. ■

The result of Theorem 13 is quite surprising, namely, that for the G-IC-OR we can achieve the gDoF region of the classical G-IC in all regimes. This is especially surprising in the strong and very strong interference regimes where joint decoding of intended and interfering messages is optimal for the classical G-IC—recall that joint decoding appears to be precluded by the absence of codebook knowledge in the G-IC-OR. This seems to suggest that while decoding of the undesired messages is not possible, one may still estimate the codewords corresponding to the undesired messages.

VIII. Finite SNR performance

In the previous section we showed that the gDoF region of the classical G-IC can be achieved even when one receiver lacks knowledge of the interfering codebook. One may then ask whether it is possible to achieve the capacity, possibly up to a constant gap, of the classical G-IC at all finite SNRs. We next show that this is indeed possible. For future use, the capacity region of



(a) Very strong interference.

(b) Strong interference.

(c) Moderately weak interference 1.

(d) Moderately weak interference 2.

Fig. 4. How to achieve the gDoF region for the G-IC-OR in different parameter regimes.

the classical G-IC is outer bounded by [17]

$$\mathcal{R}_{\text{out}}^{(\text{G-IC})}: \quad R_1 \leq \mathsf{I_g}\left(\mathsf{SNR}\right), \tag{29a}$$

$$R_2 \leq \mathsf{I_g}\left(\mathsf{SNR}\right), \tag{29b}$$

$$R_1 + R_2 \leq \left[\mathsf{I_g}\left(\mathsf{SNR}\right) - \mathsf{I_g}\left(\mathsf{INR}\right)\right]^+ + \mathsf{I_g}\left(\mathsf{SNR}+\mathsf{INR}\right), \tag{29c}$$

$$R_1 + R_2 \leq 2\mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \tag{29d}$$

$$2R_1 + R_2 \le \left[\mathsf{I_g}\left(\mathsf{SNR}\right) - \mathsf{I_g}\left(\mathsf{INR}\right)\right]^+ + \mathsf{I_g}\left(\mathsf{SNR}+\mathsf{INR}\right) + \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \quad (29e)$$

$$R_1 + 2R_2 \le \left[\mathsf{I_g}\left(\mathsf{SNR}\right) - \mathsf{I_g}\left(\mathsf{INR}\right)\right]^+ + \mathsf{I_g}\left(\mathsf{SNR}+\mathsf{INR}\right) + \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \quad (29f)$$

which is tight for $\mathsf{SNR} \le \mathsf{INR}$ and optimal to with $1/2$ bit (per channel use per user) otherwise.

The main result of this section is:

Theorem 14. *For the G-IC-OR it is possible to achieve the outer bound region in* (29) *to within* $\frac{1}{2}\log\left(8\pi\mathrm{e}\right) \approx 3.04$ *bits per channel use per user.*

Proof of Theorem 14: We consider different regimes separately.

a) Very strong interference $\mathsf{SNR}(1+\mathsf{SNR}) \le \mathsf{INR}$*:* In the regime the capacity region of the classical G-IC is given by (29a) and (29b). For achievability we consider the achievable region in Theorem 9 with

$$N = \mathsf{N_d}(\mathsf{SNR}) \quad \text{(equivalent of } \beta = 1\text{)}$$
$$\Longrightarrow N^2 - 1 \le \mathsf{SNR} \le \frac{\mathsf{INR}}{1+\mathsf{SNR}} \le \mathsf{INR}. \quad (30)$$

Recall that the achievable region in Theorem 9 is the region in (12) with the inputs as in (21); the sum-rate in Theorem 9 is redundant if $I(X_1;Y_1|X_2) + I(X_2;Y_2) \le I(X_1,X_2;Y_1)$, that is, if $I(X_2;Y_2) \le I(X_2;Y_1)$, for all input distributions in (21). With a Gaussian X_2 as in (21):

$$I(X_2;Y_2) \le I(X_2;Y_2|X_1) = I(X_{2G};\sqrt{\mathsf{SNR}}\ X_{2G} + Z_2) = \mathsf{I_g}(\mathsf{SNR}),$$

and

$$I(X_2;Y_1) = I(X_{2G};\sqrt{\mathsf{INR}}\ X_{2G} + \sqrt{\mathsf{SNR}}\ X_{1D} + Z_2) \ge \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right),$$

because a Gaussian noise is the worst noise for a Gaussian input. Since in very strong interference we have $\mathsf{I_g}(\mathsf{SNR}) \le \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right)$, the condition $I(X_2;Y_2) \le I(X_2;Y_1)$ is verified for all inputs in (21) and hence we can drop the sum-rate constraint in (22c) from Theorem 9. Therefore, in this regime the following rates are achievable

$$\mathcal{R}_{\mathrm{in}}^{\text{(G-IC-OR very strong)}}: \quad R_1 \le \mathsf{I_g}(\mathsf{SNR}) - \Delta_1, \quad (31a)$$

$$R_2 \le \mathsf{I_g}(\mathsf{SNR}) - \Delta_2, \quad (31b)$$

where

$$\Delta_1 := \mathsf{I_g}(\mathsf{SNR}) - \mathsf{I_d}\left(N, \mathsf{SNR}\right)$$
$$\leq \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right) \text{ for the reasoning leading to eq.(20),} \tag{31c}$$
$$\Delta_2 := \mathsf{I_g}\left(\min\left(N^2-1, \mathsf{INR}\right)\right) - \mathsf{I_d}\left(N, \frac{\mathsf{INR}}{1+\mathsf{SNR}}\right)$$
$$= \log(N) - \left[\log(N) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+ \leq \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right), \tag{31d}$$

where the equality in (31d) is a consequence of the relationships in (30).

It is immediate to see that (31c) is the gap for R_1 and that (31d) is the gap for R_2. Therefore in this regime the gap is at most $\frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right)$ per channel use per user, and it is due to shaping loss and integer penalty.

b) Strong interference $\mathsf{SNR} \leq \mathsf{INR} < \mathsf{SNR}(1+\mathsf{SNR})$*:* In this regime the capacity region of the classical G-IC is given by (29a)-(29c), and has two dominant corner points

$$\mathcal{R}_{\text{out}}^{\text{(G-IC strong P1)}}: \quad (R_1, R_2) = \left(\mathsf{I_g}\left(\mathsf{SNR}\right), \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right)\right), \tag{32a}$$

and

$$\mathcal{R}_{\text{out}}^{\text{(G-IC strong P2)}}: \quad (R_1, R_2) = \left(\mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right), \mathsf{I_g}\left(\mathsf{SNR}\right)\right). \tag{32b}$$

The other two corner points are $(R_1, R_2) = (\mathsf{I_g}(\mathsf{SNR}), 0)$ and $(R_1, R_2) = (0, \mathsf{I_g}(\mathsf{SNR}))$ that can be exactly achieved by silencing one of the users.

For achievability we mimic the proof of the gDoF region in the same regime (see Fig. 4(b)), that is, we show the achievability to within a constant gap of the corner points in (32a) and (32b) by choosing two different values of N in Theorem 9. For the corner point in (32a) we consider the achievable region in Theorem 9 with

$$N = \mathsf{N_d}(\mathsf{SNR}) \quad \text{(equivalent of } \beta = 1\text{)}$$
$$\Longrightarrow N^2 - 1 \leq \mathsf{SNR} \leq \mathsf{INR} \leq \mathsf{SNR}(1+\mathsf{SNR}), \tag{33a}$$

and for the corner point (32b) we consider the achievable region in Theorem 9 with

$$N = \mathsf{N_d}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right) \quad \text{(equivalent of } \beta = \alpha - 1\text{)}$$
$$\Longrightarrow N^2 - 1 \leq \frac{\mathsf{INR}}{1+\mathsf{SNR}} \leq \mathsf{SNR} \leq \mathsf{INR}. \tag{33b}$$

For the choice of N in (33a) the achievable region in Theorem 9 can be written as

$$\begin{aligned}
R_1 &\leq \mathsf{I_d}\left(N, \mathsf{SNR}\right) \\
&= \left[\log\left(N\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+, \\
R_2 &\leq \mathsf{I_d}\left(N, \frac{\mathsf{INR}}{1+\mathsf{SNR}}\right) + \mathsf{I_g}\left(\mathsf{SNR}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \mathsf{INR}\right)\right) \\
&= \left[\mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+ + \mathsf{I_g}(\mathsf{SNR}) - \log(N), \\
R_1 + R_2 &\leq \mathsf{I_d}\left(N, \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) + \mathsf{I_g}\left(\mathsf{INR}\right) \\
&= \left[\mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+ + \mathsf{I_g}(\mathsf{INR}),
\end{aligned}$$

which can further be lower bounded as

$$\begin{aligned}
\mathcal{R}_{\text{in}}^{\text{(G-IC strong P1)}} \; : \; R_1 &\leq \log\left(N\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \mathsf{I_g}\left(\mathsf{SNR}\right) - \Delta_1, && \text{(34a)} \\
R_2 &\leq \mathsf{I_g}\left(\mathsf{SNR}+\mathsf{INR}\right) - \log(N) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right) - \Delta_2, && \text{(34b)} \\
R_1 + R_2 &\leq \mathsf{I_g}\left(\mathsf{SNR}+\mathsf{INR}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \left(\mathsf{I_g}\left(\mathsf{SNR}\right) - \Delta_1\right) + \left(\mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right) - \Delta_2\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right), && \text{(34c)}
\end{aligned}$$

where the sum-rate bound is clearly redundant and where

$$\Delta_1 := \mathsf{I_g}\left(\mathsf{SNR}\right) - \log(N) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \leq \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right), \tag{35a}$$

$$\Delta_2 := \log(N) - \mathsf{I_g}\left(\mathsf{SNR}\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \leq \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right). \tag{35b}$$

Therefore, with N as in (33a) in Theorem 9, the gap to the corner point in (32a) is at most $\frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right)$ per channel use per user, as for the very strong interference regime.

By following similar steps, for the choice of N in (33b) in Theorem 9, the gap to the corner point in (32b) is still given by (35), that is, the gap is at most $\frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right)$ per channel use per user, as for the very strong interference regime.

c) Moderately weak interference $\mathsf{INR} \le \mathsf{SNR} \le \mathsf{INR}(1+\mathsf{INR})$*:* In this regime the capacity of the G-IC is outer bounded by (29).

As we did for the gDoF region (see Figs. 4(c) and 4(d)), we show here that we can achieve, up to a constant gap, all dominant corner points of (29). By silencing one of the users, we can achieve $(R_1, R_2) = (\mathsf{I_g}(\mathsf{SNR}), 0)$ and $(R_1, R_2) = (0, \mathsf{I_g}(\mathsf{SNR}))$; these rate points are to within 1 bit of the corner points of (29) given by $(R_1, R_2) = (A, \mathsf{I_g}(\mathsf{SNR}))$ and $(R_1, R_2) = (\mathsf{I_g}(\mathsf{SNR}), A)$ where

$$\begin{aligned} A &:= \mathsf{I_g}(\mathsf{SNR}+\mathsf{INR}) + \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) - \mathsf{I_g}(\mathsf{SNR}) - \mathsf{I_g}(\mathsf{INR}) \\ &= \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{SNR}}\right) + \mathsf{I_g}\left(\frac{\mathsf{SNR}}{(1+\mathsf{INR})^2}\right) \\ &\le \mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+\mathsf{SNR}}\right) + \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{INR}}\right) \le 2 \cdot \frac{1}{2}\log(2) = 1. \end{aligned}$$

We therefore have to show the achievability of the remaining two corner points obtained by the intersection of the sum-rate outer bound (given by $\min$(eq.(29c),eq.(29d))) with either (29e) or (29f). For these corner points, the gDoF-optimal choices of β were $2\alpha-1$ and $1-\alpha$, which we mimic here by choosing the following values of N in the region in (41) (a simplified achievable region from Theorem 10)

$$\begin{aligned} &N = \mathsf{N_d}\left(\frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}}\right) \quad \text{(equivalent of } \beta = 2\alpha-1) \\ &\Longrightarrow N^2 - 1 \le \frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}} \le \min\left(\frac{\mathsf{INR}^2}{1+2\mathsf{INR}}, \frac{\mathsf{INR}\cdot\mathsf{SNR}}{1+\mathsf{SNR}+2\mathsf{INR}}\right), \end{aligned} \tag{36}$$

because $\mathsf{INR} \le \mathsf{SNR}$, and

$$\begin{aligned} &N = \mathsf{N_d}\left(\frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}}\right) \quad \text{(equivalent of } \beta = 1-\alpha) \\ &\Longrightarrow N^2 - 1 \le \frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}} \le \min\left(\frac{\mathsf{INR}^2}{1+2\mathsf{INR}}, \frac{\mathsf{INR}\cdot\mathsf{SNR}}{1+\mathsf{SNR}+2\mathsf{INR}}\right), \end{aligned} \tag{37}$$

because $\mathsf{SNR} \le \mathsf{INR}(1+\mathsf{INR})$. In the regime $\mathsf{INR} \le \mathsf{SNR} \le \mathsf{INR}(1+\mathsf{INR})$ we also have

$$\frac{\mathsf{INR}^2}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}} \le \frac{\mathsf{INR}^2}{(1+\mathsf{INR})^2+\mathsf{INR}} \le 1 \le N^2-1, \quad \forall N \ge 2. \tag{38}$$

With (36)-(38), and by recalling that $\mathsf{I_g}(x) - \frac{1}{2}\log(4) \le \log(\mathsf{N_d}(x)) \le \mathsf{I_g}(x),\ x \ge 0$, the region in (41) can be further lower bounded as follows[2]

$$\mathcal{R}_{\text{in}}^{\text{(G-IC-OR weak)}}: \quad R_1 \le \mathsf{I_g}(x) - \frac{1}{2}\log(4) - \frac{1}{2}\log\left(\frac{\pi e}{3}\right) + \mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right), \tag{39a}$$

$$R_2 \le \mathsf{I_g}\left(\frac{\mathsf{INR}^2}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}}\right) - \frac{1}{2}\log\left(\frac{\pi e}{3}\right) + \mathsf{I_g}\left(\frac{\mathsf{SNR}}{2}\right) - \mathsf{I_g}(x), \tag{39b}$$

$$R_1 + R_2 \le \mathsf{I_g}\left(\min\left(\frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}}, \frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}}\right)\right) - \frac{1}{2}\log(4) + \mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) - \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{INR}}\right) + \mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right) - 2\cdot\frac{1}{2}\log\left(\frac{\pi e}{3}\right), \tag{39c}$$

where

$$x := \frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}} \text{ if } N \text{ as in (36), or} \tag{39d}$$

$$x := \frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}} \text{ if } N \text{ as in (37).} \tag{39e}$$

In Appendix B we show that region in (39) achieves the classical G-IC outer bound to within $\frac{1}{2}\log(8\pi e) \approx 3.04$ bits (per channel user per user).

d) Noisy interference $\mathsf{INR}(1+\mathsf{INR}) \le \mathsf{SNR}$*:* In this regime Gaussian inputs, treating interference as noise, and power control is optimal to within $1/2$ bit (per channel use per user) for the classical G-IC; since this scheme does not require codebook knowledge / joint decoding, the gap result applies to the G-IC-OR as well.

This conclues the proof. ■

[2] In order to get the sum-rate, let $n = N^2 - 1 \in \mathbb{N}$ and consider either $N = \mathsf{N_d}(a) : n_a := \mathsf{N_d}(a)^2 - 1 \le a \in \mathbb{R}^+$ or $N = \mathsf{N_d}(b) : n_b := \mathsf{N_d}(a)^2 - 1 \le b \in \mathbb{R}^+$ in the expression $y(n) := \mathsf{I_g}(\min(n,a)) + \mathsf{I_g}(\min(n,b)) - \mathsf{I_g}(n)$ that appears in the sum-rate. It follows easily that for $N = \mathsf{N_d}(a) \ : \ y = \mathsf{I_g}(\min(n_a, b)) \ge \mathsf{I_g}(\min(n_a, n_b)) \ge \mathsf{I_g}(\min(a,b)) - \frac{1}{2}\log(4)$, and for $N = \mathsf{N_d}(b) \ : \ y = \mathsf{I_g}(\min(a, n_b)) \ge \mathsf{I_g}(\min(n_a, n_b)) \ge \mathsf{I_g}(\min(a,b)) - \frac{1}{2}\log(4)$, where the term $\frac{1}{2}\log(4)$ is due to the "integer penalty".

IX. Conclusion

In this paper we derived capacity results for the interference channel where one of the receivers lacks knowledge of the interfering codebook, in contrast to a classical model where both receivers possess full codebook knowledge. For the class of injective semi-deterministic interference channels with one oblivious receiver, we derived a capacity result to within a constant gap; the gap is zero for fully deterministic channels, thereby providing an exact capacity characterization. We also derived the exact capacity region for a general memoryless interference channel with one oblivious receiver in the regime where the non-oblivious receiver experiences very strong interference.

We next proceeded to the Gaussian noise channel, where, unlike past work on oblivious receivers, we were able to demonstrate performance guarantees. For the symmetric case we derived the gDoF region and the capacity region to within a constant gap of $\frac{1}{2}\log(8\pi \mathrm{e}) \approx 3.04$ bits (per channel use per user). Surprisingly, this lack of codebook knowledge at one receiver does not impact the gDoF at all, and only the Gaussian capacity region to within a constant gap, compared to having full codebook knowledge. We believe this is because even though the mapping from codewords to messages may not be known, this does not prevent the receiver from estimating and removing the effect of the interfering codeword itself.

Appendix A
Proof of Theorem 10

We proceed to evaluate the rate region in Proposition 3 with the inputs in (23). With the chosen inputs, the outputs are

$$Y_1 = h_{11}\sqrt{1-\delta_1}X_{1D} + h_{11}\sqrt{\delta_1}X_{1G} + h_{12}\sqrt{1-\delta_2}X_{2Gc} + h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1,$$
$$Y_2 = h_{21}\sqrt{1-\delta_1}X_{1D} + h_{21}\sqrt{\delta_1}X_{1G} + h_{22}\sqrt{1-\delta_2}X_{2Gc} + h_{22}\sqrt{\delta_2}X_{2Gp} + Z_2.$$

The achievable region in (9) with $Q=\emptyset, U_2 = X_{2Gc}$ reduces to

$$\begin{aligned}R_1 &\le I(X_1;Y_1|X_{2Gc})\\ &= h(Y_1|X_{2Gc}) - h(Y_1|X_1,X_{2Gc})\\ &= h(h_{11}\sqrt{1-\delta_1}X_{1D} + h_{11}\sqrt{\delta_1}X_{1G} + h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1)\\ &- h(h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1)\end{aligned}$$

$$= h\left(\sqrt{\frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2\delta_2}}X_{1D}+Z_1\right)-h(Z_1)$$
$$+\mathsf{I_g}\left(|h_{11}|^2\delta_1+|h_{12}|^2\delta_2\right)-\mathsf{I_g}\left(|h_{12}|^2\delta_2\right);$$

therefore, by Theorem 8, we can further lower bound the rate of user 1 as

$$R_1 \le \mathsf{I_d}\left(N, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2\delta_2}\right)+\mathsf{I_g}\left(\frac{|h_{11}|^2\delta_1}{1+|h_{12}|^2\delta_2}\right),$$

thus proving (24a).

For the rate of user 2 we have

$$R_2 \le I(X_2;Y_2)$$
$$= h\left(h_{21}\sqrt{1-\delta_1}X_{1D}+h_{21}\sqrt{\delta_1}X_{1G}+h_{22}\sqrt{1-\delta_2}X_{2Gc}+h_{22}\sqrt{\delta_2}X_{2Gp}+Z_2\right)$$
$$-h\left(h_{21}\sqrt{1-\delta_1}X_{1D}+h_{21}\sqrt{\delta_1}X_{1G}+Z_2\right)$$
$$= h\left(\sqrt{\frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2}}X_{1D}+Z_2\right)-h(Z_2)+\mathsf{I_g}\left(|h_{21}|^2\delta_1+|h_{22}|^2\right)$$
$$-h\left(\sqrt{\frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}}X_{1D}+Z_2\right)+h(Z_2)-\mathsf{I_g}\left(|h_{21}|^2\delta_1\right)$$

therefore, by Theorem 8, we can further lower bound the rate of user 2 as

$$R_2 \le \mathsf{I_d}\left(N, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2}\right)+\mathsf{I_g}\left(\frac{|h_{22}|^2}{1+|h_{21}|^2\delta_1}\right)$$
$$-\mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right)$$

thus proving (24b).

Finally for the sum-rate we have

$$R_1+R_2 \le I(X_1,X_{2Gc};Y_1)+I(X_2;Y_2|X_{2Gc})$$
$$= h(h_{11}\sqrt{1-\delta_1}X_{1D}+h_{11}\sqrt{\delta_1}X_{1G}+h_{12}\sqrt{1-\delta_2}X_{2Gc}+h_{12}\sqrt{\delta_2}X_{2Gp}+Z_1)$$
$$-h(h_{12}\sqrt{\delta_2}X_{2Gp}+Z_1)$$
$$+h(h_{21}\sqrt{1-\delta_1}X_{1D}+h_{21}\sqrt{\delta_1}X_{1G}+h_{22}\sqrt{\delta_2}X_{2Gp}+Z_2)$$
$$-h(h_{21}\sqrt{1-\delta_1}X_{1D}+h_{21}\sqrt{\delta_1}X_{1G}+Z_2)'$$

therefore, by Theorem 8, we can further lower bound the sum-rate as

$$
\begin{aligned}
R_1 + R_2 \le\ & \mathsf{I_d}\left(N, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2}\right) + \mathsf{I_g}\left(|h_{11}|^2\delta_1+|h_{12}|^2\right) \\
& - \mathsf{I_g}\left(|h_{12}|^2\delta_2\right) \\
& + \mathsf{I_d}\left(N, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2\delta_2}\right) + \mathsf{I_g}\left(|h_{21}|^2\delta_1+|h_{22}|^2\delta_2\right) \\
& - \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right) - \mathsf{I_g}\left(|h_{21}|^2\delta_1\right)
\end{aligned}
$$

thus proving (24c).

Remark 4. *For future use, we specialized the derived achievable rate region for the power splits* $\delta_1 = \frac{1}{1+|h_{21}|^2}$ *and* $\delta_2 = \frac{1}{1+|h_{12}|^2}$ *inspired by [17]; we thus have that the following region is achievable for any* $N \in \mathbb{N}$

$$
R_1 \le \mathsf{I_d}\left(N, \frac{|h_{11}|^2 a}{1+\frac{|h_{11}|^2}{1+|h_{21}|^2}+b}\right) + \mathsf{I_g}\left(\frac{\frac{|h_{11}|^2}{1+|h_{21}|^2}}{1+b}\right), \tag{40a}
$$

$$
\begin{aligned}
R_2 \le\ & \mathsf{I_d}\left(N, \frac{|h_{21}|^2 a}{1+a+|h_{22}|^2}\right) + \mathsf{I_g}\left(\frac{|h_{22}|^2}{1+a}\right) \\
& - \mathsf{I_g}\left(\min\left(N^2-1, |h_{21}|^2 a\right)\right),
\end{aligned} \tag{40b}
$$

$$
\begin{aligned}
R_1 + R_2 \le\ & \mathsf{I_d}\left(N, \frac{|h_{11}|^2 a}{1+\frac{|h_{11}|^2}{1+|h_{21}|^2}+|h_{12}|^2}\right) + \mathsf{I_g}\left(\frac{|h_{11}|^2}{1+|h_{21}|^2}+|h_{12}|^2\right) - \mathsf{I_g}(b) \\
& + \mathsf{I_d}\left(N, \frac{|h_{21}|^2 a}{1+a+\frac{|h_{22}|^2}{1+|h_{12}|^2}}\right) + \mathsf{I_g}\left(\frac{\frac{|h_{22}|^2}{1+|h_{12}|^2}}{1+a}\right) \\
& - \mathsf{I_g}\left(\min\left(N^2-1, |h_{21}|^2 a\right)\right).
\end{aligned} \tag{40c}
$$

where $a := \frac{|h_{21}|^2}{1+|h_{21}|^2} \in [0,1]$ *and* $b := \frac{|h_{12}|^2}{1+|h_{12}|^2} \in [0,1]$.

In the symmetric case the region in (40) *is further lower bounded by*

$$
R_1 \le \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{SNR}\cdot\mathsf{INR}}{1+\mathsf{SNR}+2\mathsf{INR}}\right)\right) - \frac{1}{2}\log\left(\frac{\pi e}{3}\right) + \mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right), \tag{41a}
$$

$$
\begin{aligned}
R_2 \le\ & \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{INR}^2}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}}\right)\right) - \frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
& + \mathsf{I_g}\left(\mathsf{SNR}\frac{1}{2}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{INR}^2}{1+2\mathsf{INR}}\right)\right),
\end{aligned} \tag{41b}
$$

$$
R_1 + R_2 \le \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}}\right)\right) - \frac{1}{2}\log\left(\frac{\pi e}{3}\right)
$$

$$
\begin{aligned}
&+ \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) - \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{INR}}\right) \\
&+ \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&+ \mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{INR}^2}{1+2\mathsf{INR}}\right)\right). \qquad (41c)
\end{aligned}
$$

APPENDIX B

GAP DERIVATION FOR THE MODERATELY WEAK INTERFERENCE REGIME

In order to show achievability to within a constant gap of the outer bound in (29) by means of the achievable region in (39) (a further lower bound to the region in (41)), we distinguish two cases.

CASE 1 (regime corresponding to $\alpha \in [2/3, 1]$ in Fig. 4(c))

Assume that the sum-rate in eq.(39c) is redundant; under this condition we match the corner point of the rectangular achievable region, given by $(R_1, R_2) =$ (eq.(39a), eq.(39b)), to

$$
\begin{aligned}
\mathcal{R}_{\text{out}}^{\text{(G-IC mod P1)}}: \quad & R_1 = \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \qquad (42a)\\
& R_2 = \mathsf{I_g}(\mathsf{SNR}) - \mathsf{I_g}(\mathsf{INR}) + \mathsf{I_g}(\mathsf{INR}+\mathsf{SNR}) - \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \qquad (42b)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{R}_{\text{out}}^{\text{(G-IC mod P2)}}: \quad & R_1 = \mathsf{I_g}(\mathsf{SNR}) - \mathsf{I_g}(\mathsf{INR}) + \mathsf{I_g}(\mathsf{INR}+\mathsf{SNR}) - \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \qquad (43a)\\
& R_2 = \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \qquad (43b)
\end{aligned}
$$

which were obtained from the intersection of the sum-rate outer bound in (29c) with either (29e) or (29f). In particular, for the corner point in (42) we use x in (39d) (which corresponds to N in (36)), and for the corner point in (43) we use x in (39e) (which corresponds to N in (37)).

The gap is readily computed as follows: for the corner point in (42) we have

$$\begin{aligned}\Delta_1 &= \text{eq.(42a)-eq.(39a)}|_{x \text{ in (39d)}}\\ &\le \mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}}\right)\\ &+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)\\ &\le \frac{1}{2}\log(2)+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log\left(\frac{8\pi e}{3}\right),\end{aligned}$$

and

$$\begin{aligned}\Delta_2 &= \text{eq.(42b)-eq.(39b)}|_{x \text{ in (39d)}}\\ &\le \mathsf{I_g}(\mathsf{SNR})-\mathsf{I_g}(\mathsf{INR})+\mathsf{I_g}(\mathsf{INR}+\mathsf{SNR})-\mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)\\ &-\mathsf{I_g}\left(\frac{\mathsf{SNR}}{2}\right)+\mathsf{I_g}\left(\frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{INR}^2}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}}\right)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)\\ &\le \frac{1}{2}\log(2)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log\left(\frac{2\pi e}{3}\right),\ \text{since } \mathsf{INR}\le\mathsf{SNR} \text{ in weak interfernce;}\end{aligned}$$

while for the corner point in (43) we have

$$\begin{aligned}\Delta_1 &= \text{eq.(43a)-eq.(39a)}|_{x \text{ in (39e)}}\\ &\le \mathsf{I_g}(\mathsf{SNR})-\mathsf{I_g}(\mathsf{INR})+\mathsf{I_g}(\mathsf{INR}+\mathsf{SNR})-\mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)\\ &-\mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}}\right)+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)\\ &\le \frac{1}{2}\log(2)+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log\left(\frac{8\pi e}{3}\right),\end{aligned}$$

and

$$\begin{aligned}\Delta_2 &= \text{eq.(43b)-eq.(39b)}|_{x \text{ in (39e)}}\\ &\le \mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{SNR}}{2}\right)+\mathsf{I_g}\left(\frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}}\right)\\ &-\mathsf{I_g}\left(\frac{\mathsf{INR}^2}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}}\right)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)\\ &\le \frac{1}{2}\log(2)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log\left(\frac{2\pi e}{3}\right),\ \text{since } \mathsf{INR}\le\mathsf{SNR} \text{ in weak interference.}\end{aligned}$$

CASE 2 (regime corresponding to $\alpha \in [1/2, 2/3]$ in Fig. 4(d))

Assume that the sum-rate in (39) is not redundant, that is after simple algebraic manipulation,

$$1 + \min\left(x|_{x \text{ in (39d)}}, x|_{x \text{ in (39e)}}\right) < \underbrace{\frac{(1+2\mathsf{INR})(1+\frac{\mathsf{SNR}}{2})}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}}}_{\in[0.8565,1] \text{ for } \mathsf{INR}\leq\mathsf{SNR}\leq\mathsf{INR}(1+\mathsf{INR})} \cdot \underbrace{\frac{(1+\mathsf{INR})(1+\mathsf{INR}+\mathsf{SNR})}{(1+\mathsf{INR})^2+\mathsf{SNR}}}_{=1+x|_{x \text{ in (39e)}}},$$

which implies

$$x|_{x \text{ in (39d)}} \leq x|_{x \text{ in (39e)}}. \quad (44)$$

Under the condition in (44) we match one of the corner point of the pentagon-shaped achievable region in (39) to

$$\mathcal{R}_{\text{out}}^{\text{(G-IC weak P1)}} : R_1 = 3\mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) - \mathsf{I_g}(\mathsf{SNR}+\mathsf{INR}) - \mathsf{I_g}(\mathsf{SNR}) + \mathsf{I_g}(\mathsf{INR}), \quad (45a)$$

$$R_2 = \mathsf{I_g}(\mathsf{SNR}) - \mathsf{I_g}(\mathsf{INR}) + \mathsf{I_g}(\mathsf{SNR}+\mathsf{INR}) - \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \quad (45b)$$

and

$$\mathcal{R}_{\text{out}}^{\text{(G-IC weak P2)}} : R_1 = \mathsf{I_g}(\mathsf{SNR}) - \mathsf{I_g}(\mathsf{INR}) + \mathsf{I_g}(\mathsf{SNR}+\mathsf{INR}) - \mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right), \quad (46a)$$

$$R_2 = 3\mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) - \mathsf{I_g}(\mathsf{SNR}+\mathsf{INR}) - \mathsf{I_g}(\mathsf{SNR}) + \mathsf{I_g}(\mathsf{INR}), \quad (46b)$$

which were obtained from the intersection of the sum-rate outer bound in (29d) with either (29e) or (29f). In particular, for the corner point in (45) we use x in (39d) (which corresponds to N in (36)), and for the corner point in (46) we use x in (39e) (which corresponds to N in (37)).

The gap is readily computed as follows: for the corner point in (45) we have

$$\begin{aligned}
\Delta_1 &= \text{eq.(45a)-}\Big(\text{eq.(39c)-eq.(39b)}\Big)|_{x \text{ in (39d)}} \\
&\leq 2\mathsf{I_g}\left(\mathsf{INR} + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) - \mathsf{I_g}(\mathsf{SNR}+\mathsf{INR}) - \mathsf{I_g}(\mathsf{SNR}) + \mathsf{I_g}(\mathsf{INR}) + \mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{INR}}\right) \\
&- \mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right) + \mathsf{I_g}\left(\frac{\mathsf{SNR}}{2}\right) + \mathsf{I_g}\left(\frac{\mathsf{INR}^2}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}}\right) \\
&- 2\mathsf{I_g}\left(\frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}}\right) + \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
&= \frac{1}{2}\log\left(\frac{\left(\frac{\mathsf{SNR}}{2}+1\right)\left(\frac{\mathsf{INR}}{\mathsf{INR}+1}+1\right)\left(\frac{\mathsf{INR}^2}{\mathsf{INR}+(\mathsf{INR}+1)(\mathsf{SNR}+1)}+1\right)(\mathsf{INR}+1)\left(\mathsf{INR}+\frac{\mathsf{SNR}}{\mathsf{INR}+1}+1\right)^2}{\left(\frac{\mathsf{INR}^2}{2\,\mathsf{INR}+\mathsf{SNR}+1}+1\right)^2\left(\frac{\mathsf{SNR}}{2\,\mathsf{INR}+1}+1\right)(\mathsf{SNR}+1)(\mathsf{INR}+\mathsf{SNR}+1)}\right)
\end{aligned}$$

$$+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)$$
$$=\frac{1}{2}\log\left(\frac{(2\,\mathsf{INR}+1)^2\left(\frac{\mathsf{SNR}}{2}+1\right)(2\,\mathsf{INR}+\mathsf{SNR}+1)}{(\mathsf{INR}+1)\,(\mathsf{SNR}+1)\,(2\,\mathsf{INR}+\mathsf{SNR}+\mathsf{INR}\,\mathsf{SNR}+1)}\right)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)$$
$$\leq\frac{1}{2}\log(6)+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log(8\pi e)$$

and

$$\Delta_2=\text{eq.(45b)-eq.(39b)}|_{x\text{ in (39d)}}$$
$$\leq \mathsf{I_g}(\mathsf{SNR})-\mathsf{I_g}(\mathsf{INR})+\mathsf{I_g}(\mathsf{SNR}+\mathsf{INR})-\mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{SNR}}{2}\right)$$
$$+\mathsf{I_g}\left(\frac{\mathsf{INR}^2}{1+\mathsf{SNR}+2\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{INR}^2}{(1+\mathsf{INR})(1+\mathsf{SNR})+\mathsf{INR}}\right)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)$$
$$=\frac{1}{2}\log\left(\frac{\left(\frac{\mathsf{INR}^2}{2\,\mathsf{INR}+\mathsf{SNR}+1}+1\right)(\mathsf{SNR}+1)\,(\mathsf{INR}+\mathsf{SNR}+1)}{\left(\frac{\mathsf{SNR}}{2}+1\right)\left(\frac{\mathsf{INR}^2}{\mathsf{INR}+(\mathsf{INR}+1)\,(\mathsf{SNR}+1)}+1\right)(\mathsf{INR}+1)\left(\mathsf{INR}+\frac{\mathsf{SNR}}{\mathsf{INR}+1}+1\right)}\right)$$
$$+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)$$
$$=\frac{1}{2}\log\left(\frac{2\,(\mathsf{SNR}+1)\,(2\,\mathsf{INR}+\mathsf{SNR}+\mathsf{INR}\,\mathsf{SNR}+1)}{(\mathsf{INR}+1)\,(\mathsf{SNR}+2)\,(2\,\mathsf{INR}+\mathsf{SNR}+1)}\right)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)$$
$$\leq\frac{1}{2}\log(2)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log\left(\frac{2\pi e}{3}\right),$$

while for the corner point in (46) we have

$$\Delta_1=\text{eq.(46a)-eq.(39a)}|_{x\text{ in (39e)}}$$
$$\leq \mathsf{I_g}(\mathsf{SNR})-\mathsf{I_g}(\mathsf{INR})+\mathsf{I_g}(\mathsf{SNR}+\mathsf{INR})-\mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)-\mathsf{I_g}\left(\frac{\mathsf{SNR}}{1+2\mathsf{INR}}\right)$$
$$-\mathsf{I_g}\left(\frac{\mathsf{SNR}\cdot\mathsf{INR}}{(1+\mathsf{INR})^2+\mathsf{SNR}}\right)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)+\frac{1}{2}\log(4)$$
$$=\frac{1}{2}\log\left(\frac{(\mathsf{SNR}+1)\,(\mathsf{INR}+\mathsf{SNR}+1)}{\left(\frac{\mathsf{SNR}}{2\,\mathsf{INR}+1}+1\right)(\mathsf{INR}+1)\left(\frac{\mathsf{INR}\,\mathsf{SNR}}{\mathsf{SNR}+(\mathsf{INR}+1)^2}+1\right)\left(\mathsf{INR}+\frac{\mathsf{SNR}}{\mathsf{INR}+1}+1\right)}\right)$$
$$+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)$$
$$=\frac{1}{2}\log\left(\frac{(2\,\mathsf{INR}+1)\,(\mathsf{SNR}+1)}{(\mathsf{INR}+1)\,(2\,\mathsf{INR}+\mathsf{SNR}+1)}\right)+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)$$
$$\leq\frac{1}{2}\log(2)+\frac{1}{2}\log(4)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log\left(\frac{4\pi e}{3}\right),$$

and

$$
\begin{aligned}
\Delta_2 &= \text{eq.(46b)-}\ \Big(\text{eq.(39c)-eq.(39a)}\Big)\Big|_{x \text{ in (39e)}} \\
&\leq 2\mathsf{I_g}\left(\mathsf{INR}+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)-\mathsf{I_g}\left(\mathsf{SNR}+\mathsf{INR}\right)-\mathsf{I_g}\left(\mathsf{SNR}\right)+\mathsf{I_g}\left(\mathsf{INR}\right)+\mathsf{I_g}\left(\frac{\mathsf{INR}}{1+\mathsf{INR}}\right) \\
&+\frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
&=\frac{1}{2}\log\left(\frac{(1+2\mathsf{INR})((1+\mathsf{INR})^2+\mathsf{SNR})}{(1+\mathsf{INR})^2(1+\mathsf{SNR})(1+\mathsf{INR}+\mathsf{SNR})}\right)+\frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
&\leq 0+\frac{1}{2}\log\left(\frac{\pi e}{3}\right)=\frac{1}{2}\log\left(\frac{\pi e}{3}\right)
\end{aligned}
$$

This concludes the proof.

REFERENCES

[1] A. Dytso, N. Devroye, and D. Tuninetti, "On the capacity of interference channels with partial codebook knowledge," in *Proc. IEEE Int. Symp. Inf. Theory*, 2013, pp. 2039–2043.

[2] A. Dytso, D. Tuninetti, and N. Devroye, "On discrete alphabets for the two-user gaussian interference channel with one receiver lacking knowledge of the interfering codebook," http://arxiv.org/pdf/1401.5536.pdf.

[3] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3008 –3023, Jul. 2008.

[4] O. Simeone, E. Erkip, and S. Shamai, "On codebook information for interference relay channels with out-of-band relaying," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2880 –2888, May 2011.

[5] Y. Tian and A. Yener, "Relaying for multiple sources in the absence of codebook information," in *Proc. ASILOMAR*, Nov. 2011, pp. 1845 –1849.

[6] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.

[7] L. Ozarow and A. Wyner, "On the capacity of the gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov 1990.

[8] Y. Wu and S. Verdu, "The impact of constellation cardinality on Gaussian channel capacity," in *Proc. Allerton Conf. Commun., Control and Comp.*, Sep. 2010.

[9] E. Calvo, J. Fonollosa, and J. Vidal, "On the totally asynchronous interference channel with single-user receivers," in *Proc. IEEE Int. Symp. Inf. Theory*, 2009, pp. 2587–2591.

[10] E. Abbe and L. Zheng, "A coordinate system for gaussian networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 721–733, 2012.

[11] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 2871 –2874.

[12] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[13] G. Bresler and D. Tse, “The two-user gaussian interference channel: A deterministic view,” *European Transactions in Telecommunications*, vol. 19, pp. 333–354, Apr. 2008.

[14] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Camrbidge University Press, 2012.

[15] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981.

[16] A. El Gamal and M. Costa, “The capacity region of a class of deterministic interference channels,” *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 343–346, Mar. 1982.

[17] R. Etkin, D. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,” *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[18] G. Forney and L.-F. Wei, “Multidimensional constellations. i. introduction, figures of merit, and generalized cross constellations,” *IEEE J. Select. Areas Commun.*, vol. 7, no. 6, pp. 877–892, Aug. 1989.

[19] U. Erez, S. Litsyn, and R. Zamir, “Lattices which are good for (almost) everything,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, Oct 2005.